RIETI Discussion Paper Series 26-E-013

AI Data Centers and Electricity Demand: Taming the energy guzzlers

THORBECKE, Willem
RIETI

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 26-E-013
February 2026

AI Data Centers and Electricity Demand: Taming the energy guzzlers ※

Willem Thorbecke*
Research Institute of Economy, Trade and Industry, Japan

Abstract

Artificial intelligence (AI) use and its energy requirements are skyrocketing. This paper finds that the market capitalizations of Amazon, Google, Meta, and Microsoft have increased by more than $500 billion above predicted values since ChatGPT was launched in 2022. Nevertheless they negotiate aggressively to lower energy costs and transfer electricity expenses to other ratepayers. Their appetite for energy is also met by burning fossil fuels including coal. This paper considers how to incentivize Big Tech companies to internalize the externalities associated with data center electricity use. It also recommends innovations that can reduce AI energy demand. These include using AI itself to save energy at data centers and in the production of batteries, steel, glass, hydrogen, ammonia, and copper.

Keywords: Artificial intelligence, Data centers, Electricity grid, Utility companies, Fossil-fired power plants

JEL classification: Q40, Q48

※This study is conducted as a part of the research at the Research Institute of Economy, Trade and Industry (RIETI). The draft of this paper was presented at the RIETI DP seminar for the paper. I thank Kyoji Fukao, Yasuo Tanabe, Eiichi Tomiura, and other colleagues for excellent comments. Any errors are my responsibility.
* Willem Thorbecke. Research Institute of Economy, Trade and Industry, Tokyo, Japan. Email: willem-thorbecke@rieti.go.jp

1. Introduction

Artificial Intelligence (AI) use is exploding. OpenAI launched the generative AI chatbot ChatGPT in November 2022. By October 2025 it had 800 million weekly users, up from 400 million in February 2025 (Singh, 2025). Monthly visits reached 5.8 billion and the number of daily queries reached 2 billion.

The four leading tech firms, Amazon, Google, Meta, Microsoft, are riding the AI juggernaut. They have weaved AI into their offerings. Amazon uses AI to enhance customer service. Google offers the Gemini AI chatbot. Meta employs AI to enhance its social media and messaging apps. Microsoft partners with OpenAI.

This paper investigates how these four leading firms have performed after ChatGPT launched the AI revolution in November 2022. The results indicate that their stock market capitalizations have increased by between 500 billion dollars more than predicted for Amazon to trillions of dollars more for Meta.

Despite these windfalls, Big Tech companies negotiate aggressively with public utilities to reduce their electricity expenses. This transfers their electricity costs to other customers. Saul et al. (2025) reported that electricity bills increase more for customers closer to data centers and fall more for customers further from them.

Data centers guzzle energy and burden the grid. Tech firms will spend almost 500 billion dollars in 2025 constructing AI data centers (Stylianou, 2025). Saul et al. (2025) reported that data centers consume 39% of the electricity used in Virginia, 33% of the energy used in Oregon, 18% of the energy used in Iowa, and 15% of the energy used in Nevada and Utah. Data center energy use will accelerate.

Soaring data center demand forces utilities to prop up gas-fired and other fossil fuel-powered plants. It also causes them to maintain their oldest and dirtiest generating plants (see, e.g., Martin and Peskoe, 2025). This increases localized pollution that is not offset by companies purchasing carbon credits elsewhere.

Executives such as former Google CEO Eric Schmidt and former Microsoft CEO Bill Gates argued that society should forget about climate targets and whole-heartedly embrace AI data centers. They claimed that AI can help address climate change, so data center pollution should be ignored (see, e.g., Hyman and Tilles, 2025, and Robinson, 2024).

This paper considers how to incentivize Big Tech to be more environmentally friendly and how to protect other ratepayers from picking up their electricity tabs. It also considers innovations that could reduce data center energy use.

The next section investigates the performance of the four leading tech firms after the introduction of ChatGPT. Section 3 recounts the energy use of AI data centers. Section 4 considers innovations to increase energy efficiency at data centers. Section 5 concludes.

2. Investigating Big Tech Stock Returns

2.1 Data and Methodology

Martin and Peskoe (2025) singled out four companies leading the search for power to run data centers. These are Amazon, Google, Meta, and Microsoft. Figures 1a through 1d show the evolution of their stock prices before and after the release of ChatGPT on 30 November 2022. The figures show that, since the appearance of ChatGPT launched the AI revolution, stock prices have increased 89.1% for Amazon, 79.8% for Google, 190.5% for Meta, and 74.3% for Microsoft.

To investigate how this performance compares to what would be expected, a five-factor model and an autoregressive integrated moving average (ARIMA) model are used to predict stock returns. Stock returns are estimated using these two models over the 17 years before the appearance of ChatGPT on 30 November 2022. The resulting regression coefficients and actual out-of-sample values of the independent variables are then used to forecast stock returns from November 2022 to September 2025. This sheds light on how these firms have fared during the AI revolution relative to what would be expected.

The five-factor model draws on Hamilton (2014). He reported that changes in the ten-year constant maturity U.S. Treasury interest rate, changes in the log of the price of copper, and changes in the log of nominal effective U.S. dollar exchange rate help explain world economic growth. Two other factors related to world growth, the return on the world stock market and the change in the log of the price of World Texas Intermediate (WTI) crude oil, are also included. These five variables should help explain returns on Big Tech companies.

The estimated equations take the form:

$$\Delta R_{i,t} = \alpha_0 + \beta_1 \Delta Ten_t + \beta_2 \Delta Copper_t + \beta_3 \Delta World_t + \beta_4 \Delta NEER_t + \beta_5 \Delta WTI_t \,, \qquad (1)$$

where $\Delta R_{i,t}$ is the daily stock return for Big Tech sector i, ΔTen_t is the change in the ten-year constant maturity Treasury interest rate, $\Delta Copper_t$ is the change in the log of copper futures prices, $\Delta World_t$ is the return on the world stock market, $\Delta NEER_t$ is the change in the log of the U.S. dollar nominal effective exchange rate, and ΔWTI_t is the change in the log of the spot price for WTI crude oil.

Data on stock returns and WTI crude oil prices come from the Refinitiv Datastream database, data on copper futures from investing.com, and the other data from the Federal Reserve Bank of St. Louis FRED database.

An ARIMA model is also used to predict stock returns. ARIMA models are useful for forecasting. They use past autoregressive and moving average terms to forecast future values of the independent variable.

For Amazon, Google, and Microsoft the sample period for the estimation extends from 3 January 2006 to 29 November 2022. This provides 4,409 observations. For Meta the data are only available beginning on 18 May 2012. This provides 2,745 observations. Stock prices are then forecasted over the 30 November 2022 to 2 September 2025 period and compared to actual stock prices.

2.2 Results

Table 1 presents the results from estimating equation (1). In all four cases in column (2), increases in the 10-year Treasury rate are associated with increases in stock returns. The relationships are statistically significant in three of the four cases. The positive coefficients reflect Hamilton's (2014) observation that Treasury rate increases are associated with increases in economic growth. In all four cases in column (3), increases in copper prices are associated with decreases in stock returns. The relationships are statistically significant in three of the four cases. Copper is a vital input to data centers, and price rises decrease their profitability. In all four cases in column (4), increases in world stock returns are associated with increases in Big Tech stock returns. The relationships are statistically significant at the 1% level in all four cases. The results for world stock returns indicate that world demand is crucial for Big Tech companies.

In all four cases in column (5), U.S. dollar appreciations are associated with increases in stock returns. Since Big Tech companies have large capital stocks in the U.S., exchange rate appreciations can increase the value of this capital and thus increase firms' market values. In no cases in column (6) are increases in WTI crude oil prices associated with increases in stock returns.

Figure 2a through 2h plot actual stock prices for the Big Tech firms and stock prices forecasted using the five-factor and ARIMA models. In every case, actual stock prices far exceed predicted stock prices over the sample period. For Amazon, at the end of the sample period actual prices were 20.5% greater than forecasted by the five-factor model and 28.2% greater than forecasted by the ARIMA model. For Google, actual prices were 34.2% greater than forecasted by the five-factor model and 38.6% greater than forecasted by the ARIMA model. For Meta, actual prices were 155.4% greater than forecasted by the five-factor model and 160.6% greater than forecasted by the ARIMA model. For Microsoft, actual prices were 24.2% greater than forecasted by the five-factor model and 39.3% greater than forecasted by the ARIMA model.

Table 2 presents data on the market capitalization of these companies and on how much more than predicted their stock prices have risen. In every case stock prices have increased by at least $500 billion more than forecasted. Not only, as indicated in Figure 1, have stock prices soared since the AI revolution began, but they have increased by hundreds of billions and even trillions of dollars more than expected.

Despite these windfalls, power-guzzling Big Tech firms act to transfer electricity costs to other customers. They also force electric utilities to maintain their oldest and dirtiest generating plants. The following sections consider how stakeholders can cause tech companies to bear their own electricity costs, to save energy, and to be more environmentally-friendly.

3. AI Data Center Electricity Demand

As the AI race accelerates, Amazon, Google, Meta, Microsoft and others are splurging on data centers. As Stylianou (2025) et al. reported, they will spend almost 500 billion dollars on data centers in 2025. McKinsey (2025) estimated that data centers will add capacity of 124 gigawatts between 2025 and 2030 to meet AI demand.

The International Energy Agency (2025) forecasted that AI energy demand will explode. Under its base case scenario, demand for electricity will rise from 460 terawatts in 2024 to 1,000 terawatts in 2030 to 1,300 terawatts by 2035. For the U.S., the International Energy Agency forecasts that AI use will cause almost half of the growth in electricity demand between 2025 and 2030.

Sam Altman (2025), CEO of OpenAI, downplayed the energy requirements of AI searches. He said that an average ChatGPT query uses 0.34 watt-hours of energy and 0.000085 gallons of water. However, he provided no details about how he calculated these numbers.

Google has documented how much energy a Gemini Apps query uses (Elsworth et. al, 2025). They examined the energy used by AI accelerators, active CPUs and DRAMs, idle machines, and the infrastructure supporting data centers. They reduced their measures of energy consumption by 73% to take account of carbon-free energy that Google purchased elsewhere. They reported results only for the median Gemini query and not the mean query because the mean would put more weight on prompts using large amounts of energy. They investigated only energy use related to AI inference and not to training AI models. They reported that the median query uses 0.24 watt-hours of energy, emits 0.03 grams of carbon dioxide equivalent, and consumes 0.26 milliliters of water.

While Elsworth et. al (2025) provided more transparency than Altman (2025), there are several concerns with their results. Society is affected by AI's total energy use and carbon footprint. By only reporting the median, they downplayed energy-intensive activities such as producing videos. Delavande et al. (2025) found that producing a single short video used 90 watt-hours (Wh) while a text clarification used 0.002 Wh, text generation used 0.047 Wh, and image generation used 2.9 Wh. The median of these four numbers is 1.47 and the mean is 23.24. Researchers focusing on robustness typically report different measures, and information for the mean Gemini prompt would be useful. By multiplying mean values by estimates of the total number of prompts per month, researchers could calculate the total energy use, carbon emissions, and water requirements per month. In addition, the energy guzzled by data centers to power AI applications generates heat and requires huge amounts of water to cool the servers and to generate electricity. Elsworth et. al's calculations did not include the water needed to generate electricity (see Calma, 2025).

Another robustness measure that Elsworth et. al (2025) failed to report is location-based energy consumption. Stylianou et al. (2025) observed that the carbon credits that a company purchases do not reduce their demand for energy from the local grid. Bryan et al. (2025) reported that, since data centers need to run continuously, they cannot rely on renewables as their primary energy source. Thus their energy needs are met by fossil fuels. Credits purchased elsewhere do not reduce the cascading demand for fossil fuels within a region. Professor Hannah Daly argued that carbon credits do not decarbonize local energy supply but merely serve Big Tech.[1]

The Federal Reserve System (2025) noted that data centers stoke energy demand. Slav (2025) reported that AI is an energy monster that threatens energy security and drives investment in fossil fuel plants. The Commonwealth of Virginia (2024, page 37) found that AI will "greatly

[1] Quoted in Stylianou et al. (2025).

increase energy demand and will require construction of new generation and transmission infrastructure beyond what would have otherwise been built." They reported that these projects threaten renewable energy goals and raise customers' utility rates.

Figure 3 shows that, since the launch of ChatGPT in November 2022, average electricity prices in the U.S. have risen more than 15%. Rapier (2025) observed that the number one reason why electricity prices in the U.S. soared in 2025 is because of energy demand for AI. Saul et al. (2025), analyzing 25,000 Locational Marginal Pricing (LMP) nodes across the U.S. grid, found that 75% of LMPs within 50 miles of data centers experienced electricity price increases between 2020 and 2025. In contrast, they reported that nodes that experienced electricity price falls tend to be located farther from data centers. Electricityrates.com (2025) reported that the costs of new power infrastructure needed to supply data centers are spread across all consumers.

Costs can be transferred from tech companies to other ratepayers because of how data center energy demand interacts with U.S. utility companies and public utility commissions (PUCs). As Martin and Peskoe (2025) documented, Big Tech companies threaten to move elsewhere when rates and terms for data centers are not to their liking. Regional utilities in turn compete with each other to attract data centers. Data centers offer utilities opportunities for profitable capital investments. Big tech firms then enter special contracts with utilities. Because of obscurity and claims of confidentiality concerning data center power needs, PUCs accept these contracts after cursory investigations. The special terms offered to tech companies imply that residential ratepayers are subsidizing data centers.

Even when data centers are charged tariffs along with other customers, there is uncertainty concerning whether they will continue using the extra capacity that is built for them in the future. This implies that other ratepayers bear the risks of paying for these capital investments.

The structure of utility companies and PUCs also presents an obstacle to transitioning away from traditional energy sources. Surging demand from data centers has forced utilities to maintain their oldest and dirtiest generating plants. For instance, Martin and Peskoe (2025) documented how the Mississippi Power Company in 2025 propped up a coal plant that it was going to retire to meet surging demand. Utility companies have an incentive to meet the extra demand not by innovating but by increasing the use of gas-fired and other fossil fuel-powered plants. As Hyman and Tilles (2025) noted, the present approach is for more electrification for AI without decarbonization.

Hyman (2025) and Martin and Peskoe (2025) argued that, rather than investigating Big Tech's energy needs and deciding what the grid should do at regulatory hearings, AI companies should be required to obtain their own power, transmission and backup. In this case PUCs would not have to confront the lack of transparency surrounding AI's energy use. If data centers bore their own costs, this would protect ratepayers from subsidizing Big Tech, spare the grid from the mushrooming data center power, and give AI companies an incentive to economize on energy use.

Another factor that would incentivize Big Tech to limit the environmental costs of AI would be to demand greater transparency concerning the associated energy consumption. The leading companies have all committed to reaching net zero. Stakeholders should hold them accountable by demanding more information about their environmental footprints. This would activate what Bhagwati (1988, page 85) labeled the Dracula Effect: "exposing evil to sunlight helps to destroy it." Peer and community pressure could then goad tech firms to be more ecologically responsible.

4. Innovations to Reduce Data Center Energy Demand

One area of innovation to save energy is at the individual die level.[2] Consider, for example, if these are arranged as a three-dimensional integrated circuit (3D-IC). This contains many individual CPUs and GPUs. Manufacturers determine the minimum voltage necessary to ensure proper functioning of the die by examining the voltage at the voltage regulator. The voltage that matters, however, is the voltage at the level of the individual transistors. If greater visibility could be obtained at the transistor level, it could provide a better measure of the minimum voltage actually needed. Because of lack of visibility at the transistor level, voltage requirements are overmargined.

Mobellus is working on measuring voltage at the transistor level and dynamic voltage changes at the nanosecond level. This would provide more accurate measures of the minimum voltage level and guard against overprovisioning. It would also enable recalibration of the minimum voltage needed as workloads, models, and other factors change in the rapidly evolving AI environment. Increasing visibility at the transistor level can produce a 20% power saving.

Another area of innovation is in the size of AI models. Sperling (2025) observed that AI data centers prioritize performance over energy savings. Varoquaux et al. (2025) explained how state of the art practice is to use the largest AI models possible to perform inference. This approach multiplies the environmental footprint. Varoquaux et al. reported that for many tasks such as medical imagining and tabular learning, smaller models outperform large models. For

[2] This paragraph and the next one draws on information obtained from Sperling (2025) and www.movellus.com/ .

other tasks they found that the gains from increasing model size are limited. They argued that AI practitioners should move away from assuming that bigger models are better and employ smaller, more focused models when possible.

Sperling (2025) noted that using more data, as larger models do, also increases the heat that needs to be dissipated. Using smaller models would not only reduce the power requirements directly but also reduce the energy needed for cooling. Sperling noted that data centers pay for power twice, first to power the servers and second to cool them. The overhead for cooling can equal 30% to 40% of the total.

Cooling is an area requiring innovation. A large data center using the local water supply for cooling can consume 1.8 billion gallons per years (Yañez-Barnuevo, 2025). This is equivalent to what a town of 50,000 people uses. A more sustainable approach involves placing dielectric fluids close to processing elements or inside of packages (Sperling, 2025). This poses challenges requiring further study, especially involving how to remove heat from 3D-IC cubic structures. Other solutions including locating data centers in cooler areas, where exposure to air can facilitate cooling, or even underwater (see Morales, 2025). More research is needed to find environmentally friendly ways to cool data centers.

According to Joule's law, the heat generated increases with the square of the electrical current.[3] To lower current, researchers are investigating higher voltage power-distribution systems. Traditionally, data center racks are powered by 48-volt (V) distribution systems. Di Paolo Emilio (2025) noted that the power requirement per rack will soon increase tenfold from 100 kilowatts (kW) to 1 megawatt (MW) of power. Watt's law implies that, while delivering 100 kW at 48 V would require 2,100 amperes (A) of current, delivering 1 MW at 48 V would require

[3] This paragraph and the next one draws on Di Paolo Emilio (2025) and Boon (2025).

21,000 A. The increase in current would exponentially increase the heat generated. It would also require 450lbs of copper, increasing resistance and multiplying the energy losses.

To save energy, NVIDIA and Texas Instruments are collaborating on 800V power distribution systems. They will use gallium nitride semiconductors because these have lower conduction losses than traditional silicon devices. The challenges are daunting. The higher voltage will require complex printed circuit boards and larger safety margins. Compact intermediate bus converters must be developed to facilitate the voltage step downs before reaching the processor. Progress on these issues is imperative to improve thermal efficiency.

Another strategy is to re-architect chips so they are adiabatic. This involves recycling energy within circuits.[4] Vaire Computing is pursuing this. When a signal transitions to a new voltage level, signal energy is dissipated and turns into heat. Vaire is seeking to recapture the energy by transforming the information in reversible ways. This way the information is not lost and does not generate heat.

Architectural changes can also be made to general purpose chips. As Moyer (2025) discussed, AI data centers typically use general purpose chips. While these are versatile, they also require more energy than purpose-built processors. If data centers could use processors that are tailored to the tasks at hand, they could save energy.

Innovative Optical and Wireless Networks (IOWN) can also slash energy use. IOWN uses optical signals instead of electrical signals. It builds on the principle that light carries data more efficiently than electricity (Eguchi, 2025). It has the potential to dramatically reduce data center energy use. NTT is working on making this technology operational.

[4] This paragraph draws on Moyer (2025).

Delavande et al. (2025) reported that producing a single short video uses 45,000 times more energy that performing a text clarification. They recommended ways to reduce the energy requirements of producing videos. These include diffusion caching and low precision inference. Caching involves storing and retrieving previously used data instead of recalculating them every time. Low precision inference uses small bit integers (e.g., 8-bit or 4-bit). These methods could reduce the energy required to create videos.

AI itself can be used to reduce energy costs. For data centers it can help devise ways to reduce heat, save energy associated with voltage step downs, choose smaller inference models, reduce energy wastage along cables in between racks and servers, and conserve water. Palladino (2025) also noted that AI can reduce energy waste in the production of batteries, steel, glass, hydrogen, ammonia, copper, and other goods. AI can thus promote sustainability.

There are thus promising avenues to reduce data center energy consumption. Some are ready for use. Stakeholders must push Big Tech to adopt these innovations. Others are still being developed. Future research is needed, including through partnerships between industry and universities (Driscoll, 2025).

5. Conclusion

This paper reports that the stock market capitalizations of Amazon, Google, Meta, Microsoft have increased by trillions of dollars more than expected during the AI revolution. They have used this windfall to invest heavily in AI data centers. These centers guzzle electricity, and other consumers pay the costs through higher electricity bills. This paper recommends that Big Tech be required to obtain its own power, transmission and backup. This would protect ratepayers from subsidizing tech companies, spare the grid from the soaring power demands of data centers, and incentivize AI companies to economize on energy use.

Mushrooming data center demand forces utilities to prop up gas-fired and other fossil fuel-powered plants. It also causes them to maintain their oldest and dirtiest generating plants (see, e.g., Martin and Peskoe, 2025). This increases localized pollution and damages the health of those living nearby.

Citizens, governments, the civil society and others should demand transparency concerning the environmental costs of the AI revolution. Big Tech companies have all committed to reaching net zero. Stakeholders should hold them accountable by demanding more information about their environmental footprint. Peer and community pressure could then goad them to be more ecologically friendly.

This paper considers innovations that could be made to reduce data center energy demand. Greater visibility at the transistor level could reduce voltage overmargining. Moving away from assuming bigger AI models are better could reduce energy requirements. Placing dielectric fluids close to processing elements or inside of packages could improve cooling efficiency. Powering data racks with 800 V distribution systems rather than 48 V systems could reduce current and the concomitant heat loss. Recycling energy within circuits by transforming information in reversible ways could reduce heat loss. Tailoring chips to the tasks at hand could save energy compared to using general purpose chips. Diffusion caching and low precision inference could economize the power requirements to generate videos.

While industry, government, universities and others pursue these innovations, they should not forget Jevon's Paradox. William Stanley Jevons observed that, as coal use in the 19th century became cheaper and more efficient, coal consumption actually increased. Data center electricity demand is exploding and multiplying CO_2 emissions. As data center energy efficiency increases, society must ensure that increased AI use does not harm the environment even more.

Table 1. The Exposure of Big Tech Firms to Macroeconomic Variables

(1)	(2)	(3)	(4)	(5)	(6)
Firm	Independent Variables				
	10-year Treasury Rate	Copper Futures Price	World Stock Return	U.S. Nominal Effective Exchange Rate	WTI Oil Price
Amazon	0.034*** (0.010)	-0.061** (0.030)	0.706*** (0.051)	0.846*** (0.170)	-0.016 (0.019)
Google	0.029*** (0.005)	-0.045** (0.018)	0.662*** (0.028)	0.769*** (0.102)	-0.010 (0.010)
Meta	0.015 (0.013)	-0.086** (0.039)	0.761*** (0.083)	0.736*** (0.235)	0.006 (0.026)
Microsoft	0.026*** (0.004)	-0.015 (0.016)	0.647*** (0.025)	0.754*** (0.088)	-0.009 (0.009)

Notes: The table presents regression coefficients from a regression of daily stock returns for the firms listed in column (1) on the change in the ten-year constant maturity U.S. Treasury interest rate (column 2), the change in the log of copper futures prices (column 3), the return on the world stock market (column 4), the change in the log of the U.S. dollar nominal effective exchange rate (column 5), and the change in the log of the spot price for WTI crude oil (column 6). The sample period extends from 3 January 2006 to 29 November 2022 except for Meta. In this case the period extends from 18 May 2012 to 29 November 2022. Heteroskedasticity- and autocorrelation-consistent standard errors are in parentheses. *** (**) denote significance at the 1% (5%) levels.

Table 2. Percentage and Dollar Value of Surpluses of Actual Big Tech Stock Values Relative to Predicted Values.

(1)	(2)	(3)	(4)	(5)	(6)
Firm	Stock Market Capitalization in Sept. 2025 (billions of USD)	Percentage Surplus of Actual Stock Prices in Sept. 2025 Relative to Predicted Stock Prices (Five-Factor Model)	Percentage Surplus of Actual Stock Prices in Sept. 2025 Relative to Predicted Stock Prices (ARIMA Model)	Dollar Value of Surplus of Actual Stock Prices in Sept. 2025 Relative to Predicted Stock Prices (Five-Factor Model) (billions of USD)	Dollar Value of Surplus of Actual Stock Prices in Sept. 2025 Relative to Predicted Stock Prices (ARIMA Model) (billions of USD)
Amazon	2,499	20.5%	28.2%	512	705
Google	3,039	34.2%	38.6%	1,039	1,173
Meta	1,921	155.4%	160.6%	2,985	3,085
Microsoft	3,783	24.2%	39.3%	916	1,487

Notes: Stock prices for the firms listed in column (3) using a five-factor model and in column (4) using an autoregressive integrated moving average (ARIMA) model. For Amazon, Google, and Microsoft, these models are estimated over the 3 January 2006 to 29 November 2022 period. For Meta, they are estimated over the 18 May 2012 to 29 November 2022 period. For the five-factor model, stock returns for the Big Tech companies are regressed on the change in the ten-year constant maturity U.S. Treasury interest rate, the change in the log of copper futures prices, the return on the world stock market, the change in the log of the U.S. dollar nominal effective exchange rate, and the change in the log of the spot price for WTI crude oil. From when Chat GPT was first introduced on 30 November 2022 until 2 September 2025, the regression coefficients from the five-factor and ARIMA models and actual out-of-sample values of the right-hand side variables are used to forecast stock returns. Column (3) presents the percentage surplus of actual over predicted values at the end of the forecasting period using the five-factor model and column (4) presents the percentage surplus using the ARIMA model. Columns (5) and (6) present the corresponding dollar values of the surpluses. Column (5) is calculated by multiplying column (2) with column (3) and column (6) by multiplying column (2) with column (4).

Figure 1a. Amazon's Stock Price

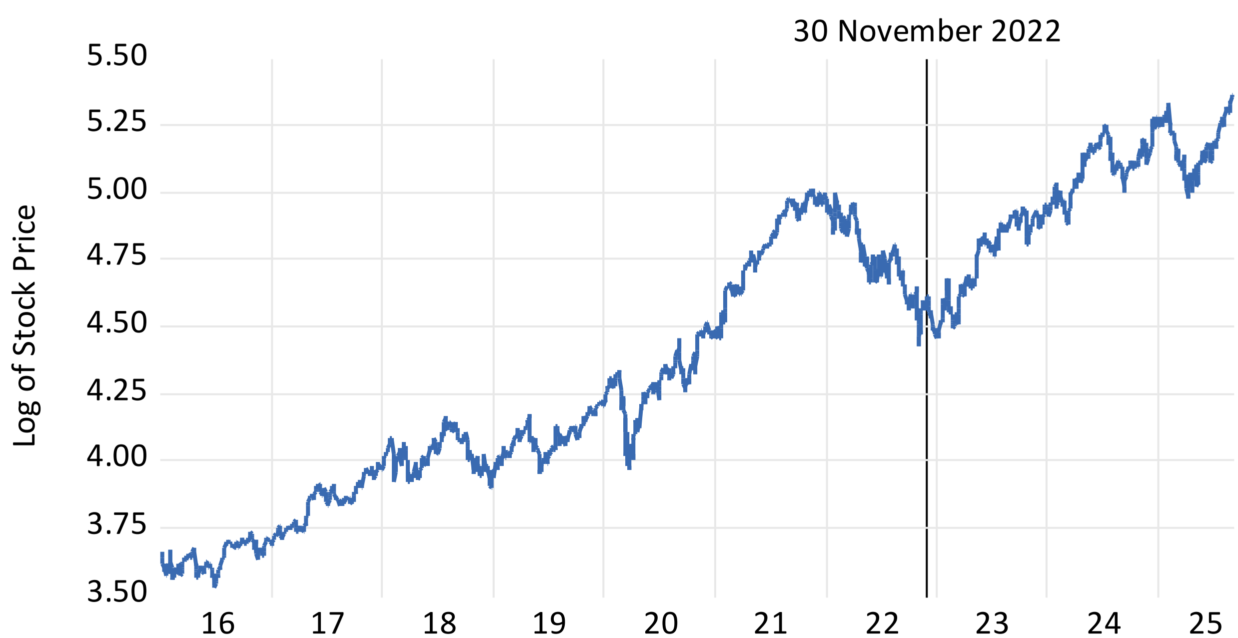

Figure 1b. Google's Stock Price

Figure 1c. Meta's Stock Price

Figure 1d. Microsoft's Stock Price
Source for Figures 1a-1d: Refinitiv Datastream database.

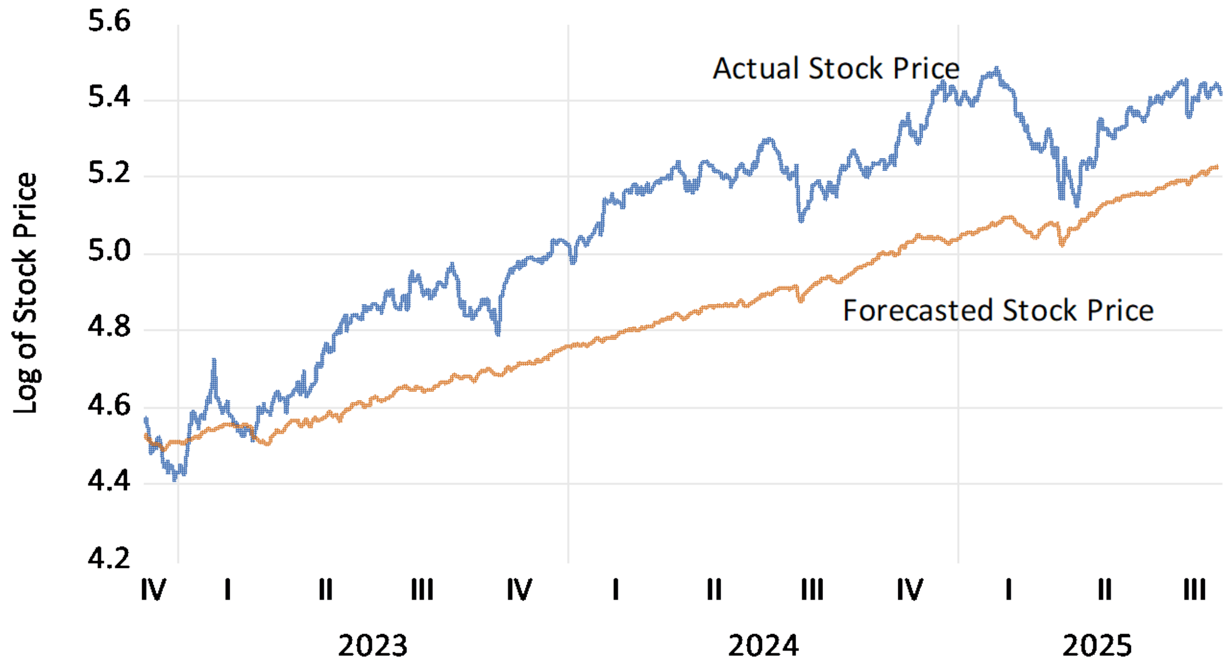

Figure 2a. Amazon's Actual Stock Prices and Prices Forecasted Using a Five-Factor Model.

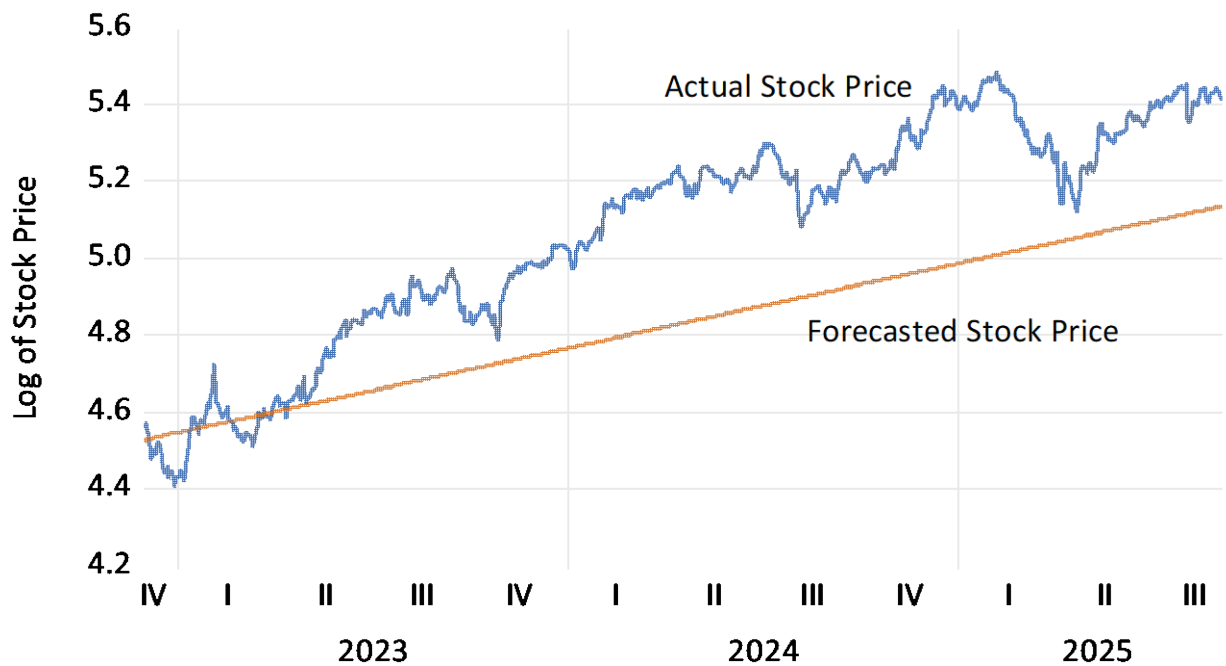

Figure 2b. Amazon's Actual Stock Prices and Prices Forecasted Using an ARIMA Model.

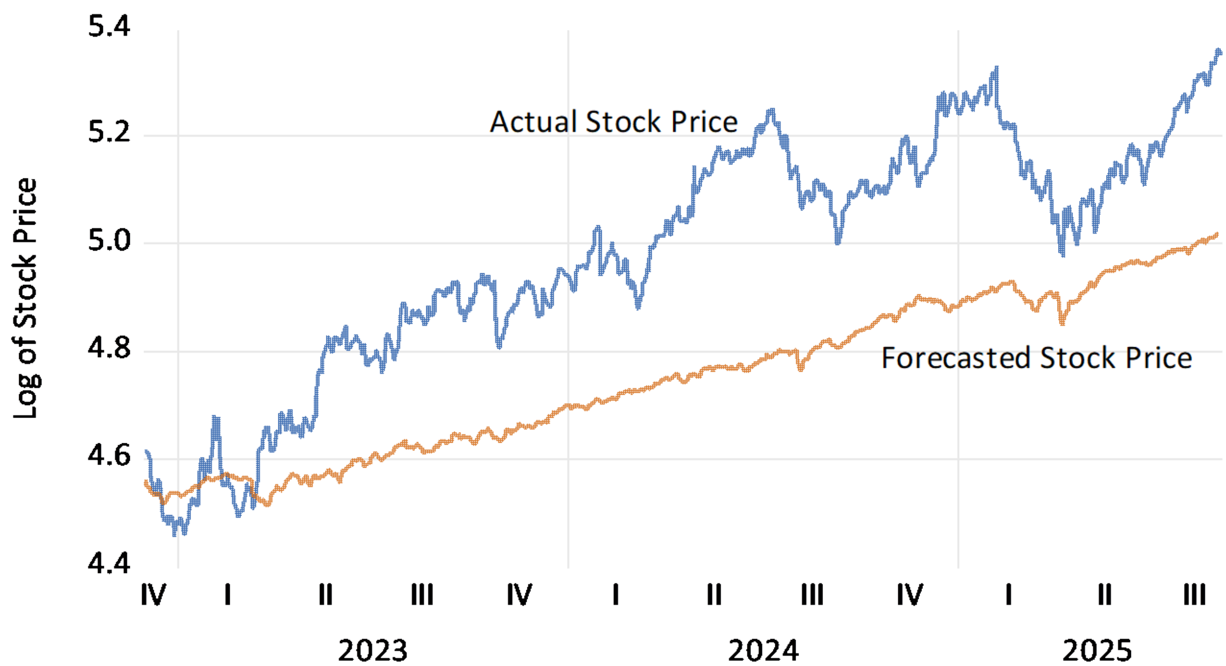

Figure 2c. Google's Actual Stock Prices and Prices Forecasted Using a Five-Factor Model.

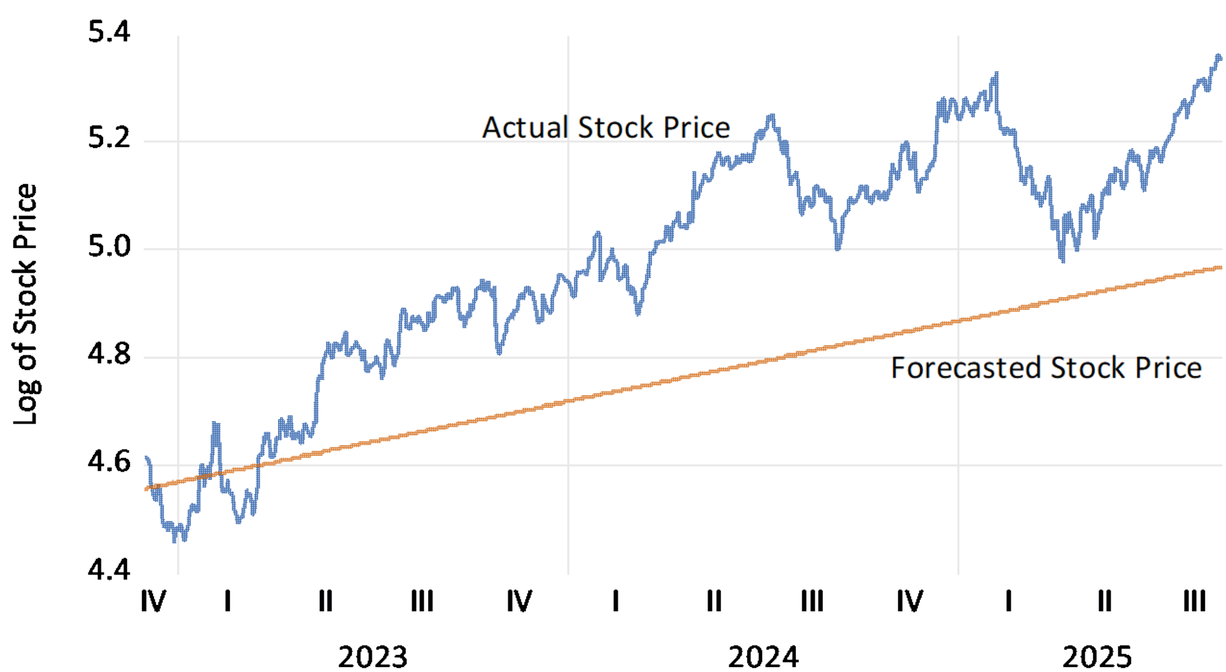

Figure 2d. Google's Actual Stock Prices and Prices Forecasted Using an ARIMA Model.

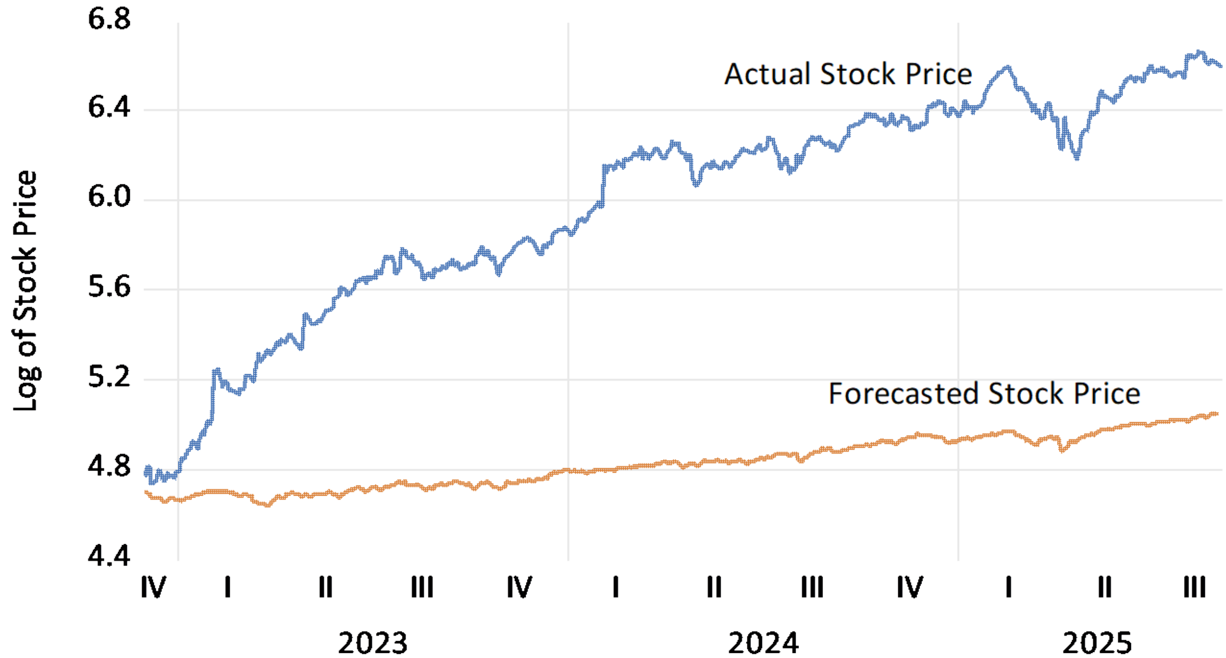

Figure 2e. Meta's Actual Stock Prices and Prices Forecasted Using a Five-Factor Model.

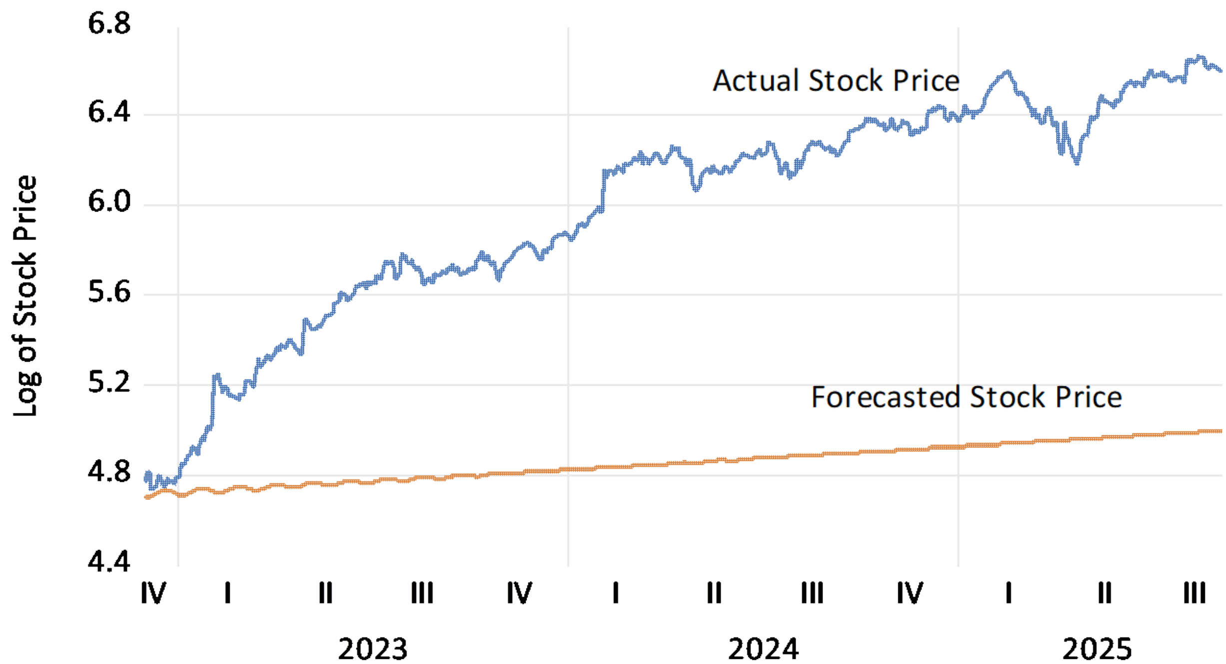

Figure 2f. Meta's Actual Stock Prices and Prices Forecasted Using an ARIMA Model.

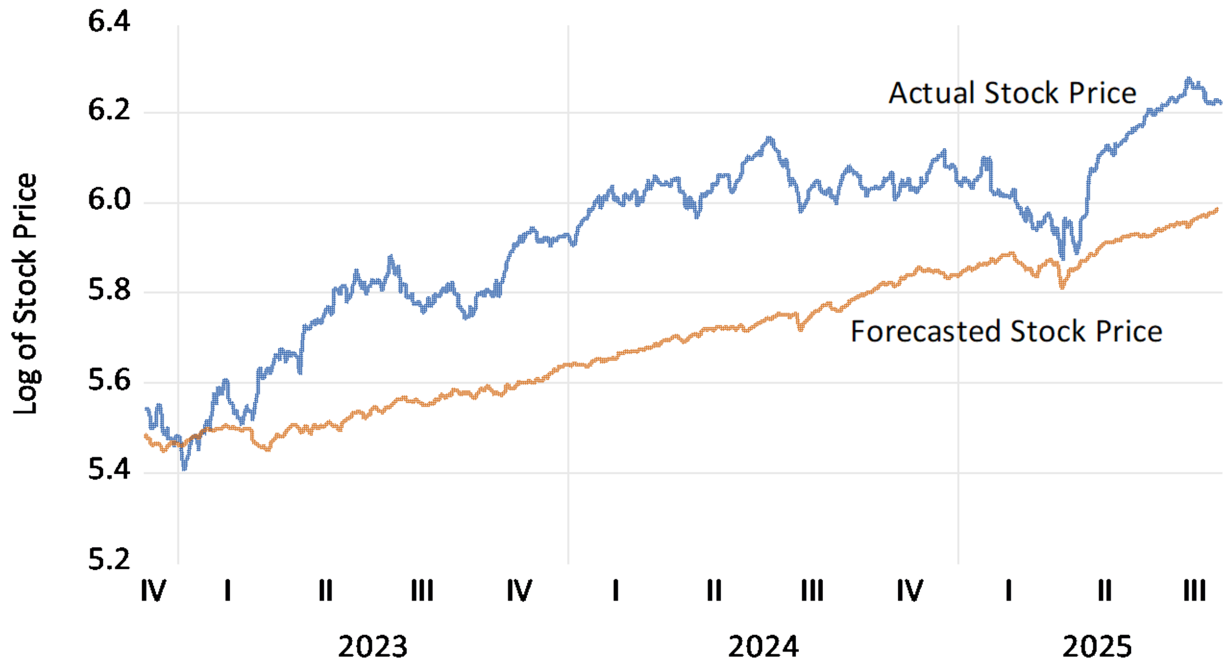

Figure 2g. Microsoft's Actual Stock Prices and Prices Forecasted Using a Five-Factor Model.

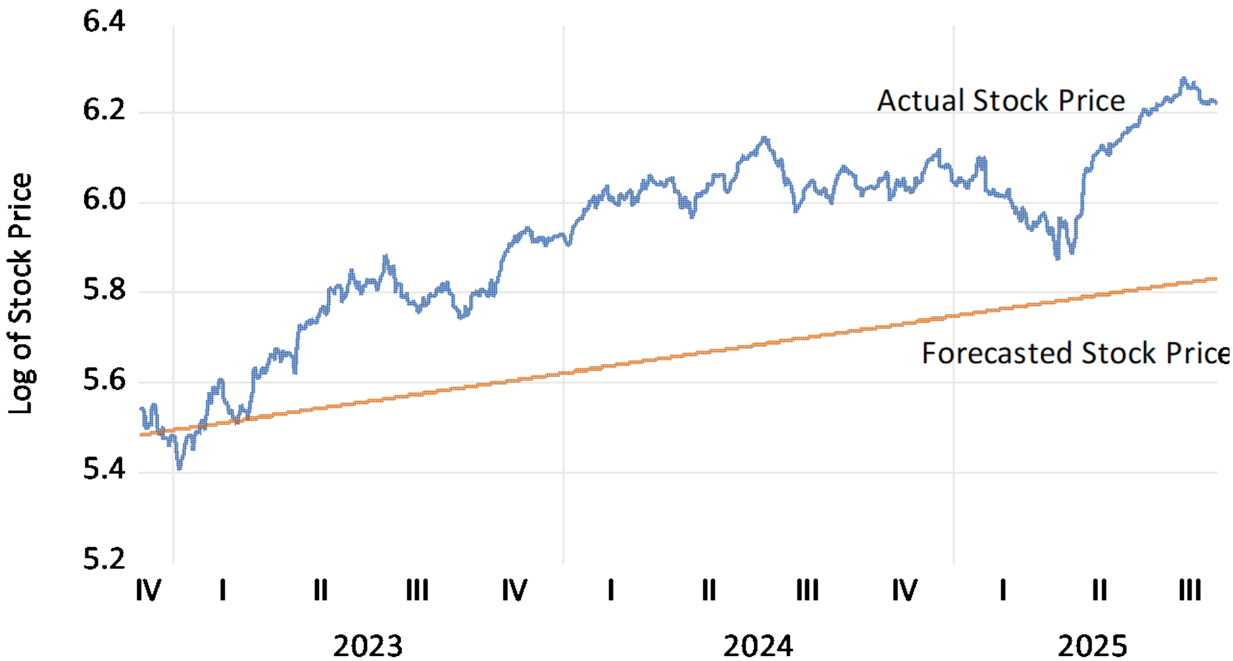

Figure 2h. Microsoft's Actual Stock Prices and Prices Forecasted Using an ARIMA Model.
Notes to Figures 2a-2h: The figures present actual stock prices for the Big Tech firms from 30 November 2022, when ChatGPT was first introduced, until 2 September 2025. Figures 2a, c, and g also present stock prices forecasted from a five-factor model estimated over the 3 January 2006 to 29 November 2022 period. Figure 2e presents stock prices forecasted from a five-factor model estimated over the 18 May 2012 to 29 November 2022 period. Actual out-of-sample values of the five factors are used to forecast returns over the 30 November 2022 to 2 September 2025 period. Figures 2b, d, and h also present stock prices forecasted from an autoregressive integrated moving average (ARIMA) model over the 3 January 2006 to 29 November 2022 period. Figure 2f also presents stock prices forecasted from an ARIMA model over the 18 May 2012 to 29 November 2022 period. The five factors used to estimate and predict stock returns are the change in the ten-year constant maturity U.S. Treasury interest rate, the change in the log of copper futures prices, the return on the world stock market, the change in the log of the U.S. dollar nominal effective exchange rate, and the change in the log of the spot price for WTI crude oil.

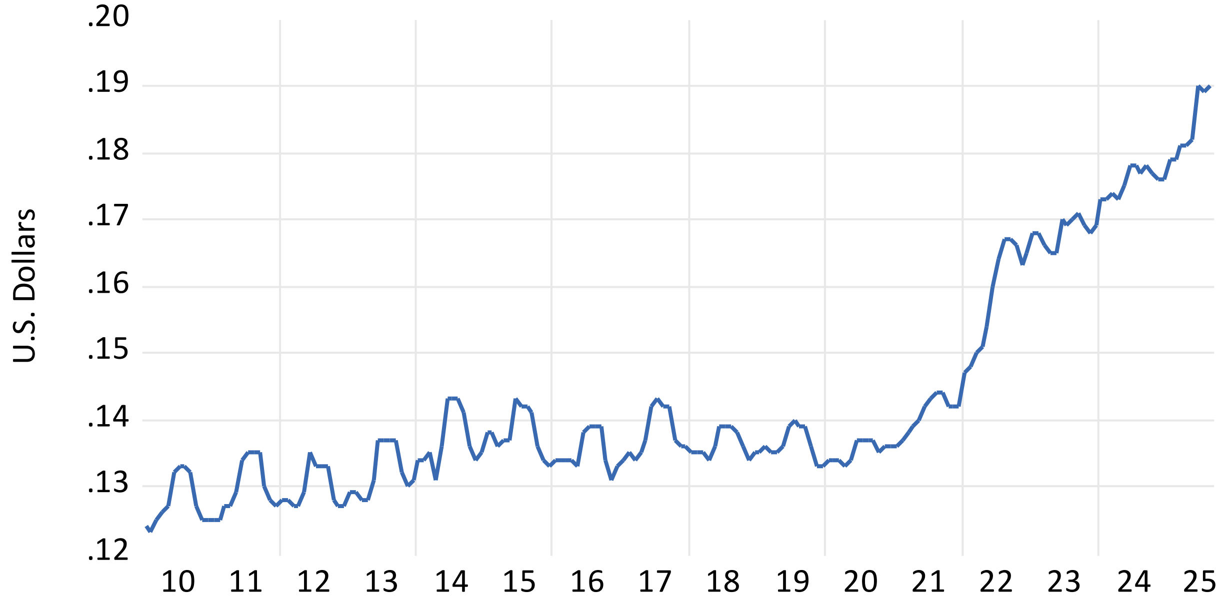

Figure 3. Average Price of Electricity per Kilowatt-Hour in U.S. Cities
Source: Federal Reserve Bank of St. Louis Fred Database.

References

Altman, S. 2025. The Gentle Singularity. Posthaven Weblog, 11 June. Available at: https://blog.samaltman.com/ .

Bhagwati, J.N. 1988. *Protectionism*. Cambridge: MIT Press.

Boon, Y.T. 2025. The 800-Volt Gorilla Inside AI Data Centers. Neuberger Berman Weblog, 1 July. Available at: https://www.nb.com/en/global/insights/the-800-volt-gorilla-inside-ai-data-centers .

Bryan, K., Cornish, C., K., Tauschinski, J. , Stylianou, N., Webber, J., Xiao, E., and Walker, O. 2025. Inside the AI Race: Can Data Centres Ever Truly Be Green? *Financial Times,* 7 August.

Calma, J. 2025. Google Says a Typical AI Text Prompt Only Uses 5 Drops of Water — Experts Say That's Misleading. *The Verge*, 21 August. Available at: https://www.theverge.com/report/763080/google-ai-gemini-water-energy-emissions-study .

Commonwealth of Virginia. 2024. Data Centers in Virginia. Report to the Governor and the General Assembly of Virginia. Available at: https://jlarc.virginia.gov/pdfs/reports/Rpt598.pdf

Delavande, J., Pierrard, R., and Luccioni, S. 2025. Video Killed the Energy Budget: Characterizing the Latency and Power Regimes of Open Text-to-Video Models. Preprint, September. Available at: https://arxiv.org/pdf/2509.19222? .

Di Paolo Emilio, M. 2025. Pioneering 800V HVDC Power Distribution for Next-Generation AI Data Centers. Power Electronic News Weblog, 31 July. Available at: https://www.powerelectronicsnews.com/pioneering-800v-hvdc-power-distribution-for-next-generation-ai-data-centers/ .

Driscoll, J. 2025. How to Avoid the Looming AI Energy Bottleneck. *Financial Times*, 8 October.

Eguchi, H. 2025. What Makes IOWN So Impressive? NTTDATA Weblog. Available at: https://www.nttdata.com/global/en/insights/focus/2025/039 .

Electricityrates.com. 2025. Rising Energy Costs From Data Centers: Who Pays the Price? Weblog, 25 August. Available at: https://electricityrates.com/resources/rising-energy-costs/ .

Elsworth, C., Huang, K., Patterson, D., Schneider, I., Sedivy, R., Goodman,S., Townsend, B., Ranganathan, P., Dean, J., Vahdat A., Gomes, B., and Manyika, J. 2025. Measuring the Environmental Impact of Delivering AI at Google Scale. Google Working Paper. Available at: https://arxiv.org/pdf/2508.15734 .

Federal Reserve System. 2025. *Beige Book*. August. Available at: https://www.federalreserve.gov/monetarypolicy/files/BeigeBook_20250903.pdf .

Hamilton, J. Oil Prices as an Indicator of Global Economic Conditions. Econbrowser Weblog, 14 December. Available at: https://econbrowser.com/archives/2014/12/oil-prices-as-an-indicator-of-global-economic-conditions .

Hyman, L. 2025. AI Firms Must Pay in Full for Energy; They Can Afford It. *Financial Times*, 26 May.

Hyman, L. and Tilles, W. 2025 The Hidden Cost of Electrification in the United States. OilPrice Weblog, 13 October. Available at: https://oilprice.com/Energy/Energy-General/The-Hidden-Cost-of-Electrification-in-the-United-States.html .

International Energy Agency. 2025. *Energy and AI.* Paris: International Energy Agency. Available at: https://iea.blob.core.windows.net/assets/601eaec9-ba91-4623-819b-4ded331ec9e8/EnergyandAI.pdf .

Martin, E., and Peskoe, A. 2025. Extracting Profits from the Public: How Utility Ratepayers Are Paying for Big Tech's Power. Harvard Law School Working Paper. Available at: https://eelp.law.harvard.edu/wp-content/uploads/2025/03/Harvard-ELI-Extracting-Profits-from-the-Public.pdf .

McKinsey. 2025. The cost of compute: A $7 trillion race to scale data centers. 28 April. Available at: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-cost-of-compute-a-7-trillion-dollar-race-to-scale-data-centers#/

Morales, J. 2025. China to Launch Commercial Underwater Data Center — Facility Expected to Consume 90% Less Power For Cooling. Tom's Hardware Weblog, 4 October. Available at: https://tech.yahoo.com/science/articles/china-launch-commercial-underwater-data-172248887.html .

Moyer, B. 2025. Will New Processor Architectures Raise Energy Efficiency? *Semiconductor Engineering*, 14 August. Available at: https://semiengineering.com/will-new-processor-architectures-raise-energy-efficiency/ .

Palladino, C. 2025. How AI Might Save More Energy than it Soaks Up. *Financial Times*, August 17.

Rapier, R. 2025. The Real Reasons Your Power Bill Is Exploding. Oilprice.com Weblog, 22 August. Available at: https://oilprice.com/Energy/Energy-General/The-Real-Reasons-Your-Power-Bill-Is-Exploding.html .

Robinson, D. 2024. Bill Gates Says not to Worry about AI Gobbling Up Energy, Tech Will Adapt. The Register, 28 June. Available at: https://www.theregister.com/2024/06/28/bill_gates_ai_power_consumption/ .

Saul, J., Nicoletti, L., Pogkas, D., Bass, D., and Malik, N. 2025. AI Data Centers Are Sending Power Bills Soaring. *Bloomberg*, 30 September.

Singh, S. 2025. Latest ChatGPT Users Stats 2025 (Growth & Usage Report). Demandsage Weblog, 7 October. https://www.demandsage.com/chatgpt-statistics/ .

Slav, I. 2025. The Energy Monster AI Is Creating. Oilprice.com Weblog, 6 September. Available at: https://oilprice.com/Energy/Energy-General/The-Energy-Monster-AI-Is-Creating.html .

Sperling, J. 2025. Crisis Ahead: Power Consumption In AI Data Centers. *Semiconductor Engineering*, 21 July. Available at https://semiengineering.com/crisis-ahead-power-consumption-in-ai-data-centers/ .

Stein, Z. 2024. Gigawatt (GW). Carbon Collective Weblog, 9 January. Available at: https://www.carboncollective.co/sustainable-investing/gigawatt-gw .

Stylianou, N., Learner, S., Bradshaw, T., Uddin, R., Bott,I., Nevitt, C., Clark, D., and Joiner, S. 2025. Inside the Relentless Race for AI Capacity. *Financial Times,* 31 July.

Varoquaux, G., Luccioni, S., and Whittaker, M. 2025. Hype, Sustainability, and the Price of the Bigger-is-Better Paradigm in AI. Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, 61-75. Available at: https://dl.acm.org/doi/pdf/10.1145/3715275.3732006 .

Yañez-Barnuevo, M. 2025. Data Centers and Water Consumption. Environmental and Energy Study Weblog, 25 June. Available at: https://www.eesi.org/articles/view/data-centers-and-water-consumption