INTERNATIONAL MONETARY FUND

Power Hungry: How AI Will Drive Energy Demand

Christian Bogmans, Patricia Gomez-Gonzalez, Ganchimeg Ganpurev, Giovanni Melina, Andrea Pescatori, Sneha Thube

WP/25/81

2025
APR

WORKING PAPER

 WP/25/81

IMF Working Paper

Research Department

Power Hungry: How AI Will Drive Energy Demand

Prepared by Christian Bogmans, Ganchimeg Ganpurev, Patricia Gomez-Gonzalez, Giovanni Melina, Andrea Pescatori, and Sneha Tube

Authorized for distribution by Petya Koeva Brooks
April 2025

ABSTRACT: The development and deployment of large language models like ChatGPT across the world requires expanding data centers that consume vast amounts of electricity. Using descriptive statistics and a multi-country computable general equilibrium model (IMF-ENV), we examine how AI-driven data center growth affects electricity consumption, electricity prices, and carbon emissions. Our analysis of national accounts reveals AI-producing sectors in the U.S. have grown nearly triple the rate of the private non-farm business sector, with firm-level evidence showing electricity costs for vertically integrated AI companies nearly doubled between 2019-2023. Simulating AI scenarios in the IMF-ENV model based on projected data center power consumption up to 2030, we find the AI boom will cause manageable but varying increases in energy prices and emissions depending on policies and infrastructure constraints. Under scenarios with constrained growth in renewable energy capacity and limited expansion of transmission infrastructure, U.S. electricity prices could increase by 8.6%, while U.S. and global carbon emissions would rise by 5.5% and 1.2% respectively under current policies. Our findings highlight the importance of aligning energy policies with AI development to support this technological revolution, while mitigating environmental impacts.

RECOMMENDED CITATION: Bogmans, Christian, Patricia Gomez-Gonzalez, Ganchimeg Ganpurev, Giovanni Melina, Pescatori, Andrea, Thube, Sneha (2025). “Power Hungry: How AI Will Drive Energy Demand?”. IMF Working Paper WP/25/81. International Monetary Fund, Washington DC.

JEL Classification Numbers:	C68, Q43, Q48, Q54, O44, L86
Keywords:	generative AI; data centers; energy and the macroeconomy; climate change and growth; CGE models.

Author’s E-Mail Address:	cbogmans@imf.org; pgomez-gonzales@imf.org; gganpurev@imf.org; gmelina@imf.org; apescatori@imf.org; sthube@imf.org

WORKING PAPERS

Power Hungry: How AI Will Drive Energy Demand

Prepared by Christian Bogmans, Ganchimeg Ganpurev, Patricia Gomez-Gonzalez, Giovanni Melina, Andrea Pescatori, and Sneha Tube[1]

[1] The authors would like to thank Petya Koeva Brooks, Domenico Giannone, Pierre-Olivier Gourinchas, Florence Jaumotte, Ryan Kellogg, Jean-Marc Natal, and Antonio Spilimbergo for their valuable suggestions. We have benefited from many discussions with and suggestions from participants of the IMF Research Department Seminar. The authors declare that they did not receive external funding for this research and that there are no conflicts of interest.

Power Hungry: How AI Will Drive Energy Demand

Christian Bogmans* Ganchimeg Ganpurev† Patricia Gomez-Gonzalez‡
Giovanni Melina§ Andrea Pescatori¶ Sneha Thube||

This version: April 10, 2025

Abstract

The development and deployment of large language models like ChatGPT across the world requires expanding data centers that consume vast amounts of electricity. Using descriptive statistics and a multi-country computable general equilibrium model (IMF-ENV), we examine how AI-driven data center growth affects electricity consumption, electricity prices, and carbon emissions. Our analysis of national accounts reveals AI-producing sectors in the U.S. have grown nearly triple the rate of the private non-farm business sector, with firm-level evidence showing electricity costs for vertically integrated AI companies nearly doubled between 2019-2023. Simulating AI scenarios in the IMF-ENV model based on projected data center power consumption up to 2030, we find the AI boom will cause manageable but varying increases in energy prices and emissions depending on policies and infrastructure constraints. Under scenarios with constrained growth in renewable energy capacity and limited expansion of transmission infrastructure, U.S. electricity prices could increase by 8.6%, while U.S. and global carbon emissions would rise by 5.5% and 1.2% respectively under current policies. Our findings highlight the importance of aligning energy policies with AI development to support this technological revolution, while mitigating environmental impacts.

JEL classification: C68, Q43, Q48, Q54, O44, L86

Keywords: generative AI; data centers; energy and the macroeconomy; climate change and growth; CGE models.

---

*Research Department, International Monetary Fund. Email: cbogmans@imf.org

†Research Department, International Monetary Fund. Email: gganpurev@imf.org

‡Research Department, International Monetary Fund. Email: pgomez-gonzalez@imf.org.

§Research Department, International Monetary Fund. Email: gmelina@imf.org.

¶Research Department, International Monetary Fund. Email: apescatori@imf.org.

||Research Department, International Monetary Fund. Email: sthube@imf.org.

The authors would like to thank Petya Koeva Brooks, Domenico Giannone, Pierre-Olivier Gourinchas, Florence Jaumotte, Ryan Kellogg, Jean-Marc Natal, and Antonio Spilimbergo for their valuable suggestions. We have benefited from many discussions with and suggestions from participants of the IMF Research Department Seminar. The authors declare that they did not receive external funding for this research and that there are no conflicts of interest.

1 Introduction

The rapid rise of generative AI in recent years often evokes images of a world becoming increasingly digital and virtual. However, the economics of AI remain firmly grounded in the physical reality of commodities, especially energy. The rapid development and adoption of large language models (LLMs), like ChatGPT, Claude, DeepSeek, and Grok, in the US, China, Europe and across the globe require the construction of an increasing number of data centers that consume vast amounts of electricity. LLM costs have two main components: a large, fixed cost for training the model on large quantities of data, and variable costs for operating and responding to user prompts (Korinek and Vipra 2024). As substantial computational resources are required during both stages, electricity consumption represents a critical input for companies delivering AI services. In Northern Virginia, which features the largest concentration of data centers in the world, the square footage of server-filled warehouses is now roughly equivalent to the floor space of 8 Empire State buildings (Cushman and Wakefield 2024, Farrell and Newman 2023).

Using descriptive statistics and a multi-country computable general equilibrium (CGE) model, IMF-ENV (Chateau et al. 2025), this paper seeks to understand how the projected growth of data centers fueled by AI will drive electricity consumption, by answering the following questions: (1) how fast have sectors involved in the development and delivery of AI-related services grown in recent years, and what has happened to the electricity consumption of leading U.S. firms in the AI production ecosystem? (2) what is the impact on energy prices and the mix of electricity sources under alternative policy scenarios? (3) what will be the impact of data centers growth on carbon emissions?

To identify AI-related economic activity in national accounts, we begin by examining where the AI production ecosystem appears in official statistics. We classify the AI production ecosystem into four firm categories: pure data center operators, AI research labs, cloud services providers, and vertically integrated companies. These firms' activities primarily align with two North American Industry Classification System (NAICS) sectors: "Data Processing, Internet Publishing, and Other Information Services" (518/519) and "Computer Systems Design and Related Services" (5415). We analyze U.S. national accounts data for these sectors to identify patterns in output growth, productivity, and growth sources (capital, intermediates including energy, and TFP). Additionally, we collect firm-level data from sustainability and annual reports to track electricity cost shares across different AI ecosystem participants.

Our analysis of national accounts and firm-level data reveals three key findings. First, AI-producing sectors in the U.S. economy have grown at nearly triple the rate of the overall private non-farm business sector. Second, this exceptional growth derives from increased total factor productivity, higher capital investment, and greater use of complementary resources, including energy. Third, firm-level evidence demonstrates that electricity costs as a share of total expenses for vertically integrated AI companies doubled between 2019 and 2023, rising from 0.8% to 1.6% on average, reflecting the rapid expansion of energy-intensive data center operations. Notably, pure data center companies maintain a substantially higher electricity cost share of 13-15%, suggesting considerable room for further increases in electricity intensity among vertically integrated AI companies as they expand their data centers.

To assess the future implications of rising AI-driven electricity demand, we employ the IMF-ENV model (Chateau et al. 2025), a multi-region, multi-sector computable general equilibrium framework. We incorporate projected data center power consumption from 2024 to 2030 across three key regions—the United States, Europe, and China—using forecasts from McKinsey and JP Morgan. These projections anticipate annual power demand growth rates of 22%, 13%, and 10% for these regions, respectively. We model AI impact by increasing IT sectors' TFP to match the anticipated growth in data center power demand. Our analysis explores three scenarios: a baseline scenario without AI-related TFP growth; an AI scenario under current energy policies; and an AI scenario with alternative energy policies aligning renewable electricity generation with regional long-term strategies. Additionally, we simulate variations accounting for potential medium-term constraints that might limit AI sector growth: one scenario caps renewable energy expansion between 2025-2030 to historical five-year average growth rates, addressing concerns about supply chain limitations, permitting delays, and policy uncertainty; the other restricts new investment in transmission and distribution infrastructure relative to baseline projections, reflecting potential challenges in grid modernization and expansion.

Our IMF-ENV model simulations reveal that the projected AI boom will cause manageable increases in energy prices and emissions compared to baseline projections, with impacts varying based on energy policies and infrastructure constraints. In the best-case scenario—with renewable energy subsidies and no infrastructure limitations—electricity prices would rise by only 0.9%. However, with constrained renewable scale-up and limited transmission infrastructure investment, prices could increase by 8.6%. Regarding emissions, we project that U.S. carbon emissions would increase by 5.5% while global emissions would rise by 1.2% on average under current policies without infrastructure constraints. The cumulative additional emissions between 2025-2030 would roughly equal Italy's entire energy-related greenhouse gas emissions over five years. Implementing policies aligned with Nationally Determined Contributions would reduce these increases by 24%. While the social cost of these additional emissions represents a small fraction of AI's expected economic benefits, they would nonetheless contribute to an already concerning accumulation of greenhouse gases.

This paper makes several contributions to the literature on AI and energy by integrating components previously studied in isolation. First, we offer a consistent quantitative evaluation of the macroeconomic impacts arising from the growing expansion of data centers within a computable general equilibrium framework, providing insights into the wider economic consequences of this structural change, particularly energy inputs. Second, we evaluate medium-term factors influencing this transition and their effects on key policy outcomes, including electricity prices, value-added changes, and international commodity prices. Finally, we examine the relationship between AI-related data center expansion and energy policies, investigating how these policies can support this phase of the digital transition while mitigating greenhouse gas emissions.

Public discourse and research in economics on generative artificial intelligence (gen-AI) has primarily focused on concerns like job displacement, safety risks, and its potential to reshape income and wealth disparities, as well as promises of boosting productivity and reinvigorating growth (Comunale and Manera 2024, Cazzaniga et al. 2024, among many others). In contrast to this literature, which largely abstracts from the physical dependencies of AI, we explicitly study the impact of projected increases in electricity demand from AI on energy prices and carbon emissions, focusing on AI production rather than adoption or task automation. Additionally, our comparative growth decomposition exercise for the U.S. economy offers novel empirical insights about the production of AI services that are absent from the existing literature.

A growing body of literature examines rising energy demand from data centers. De

Vries (2023) estimates AI-related power consumption could reach 127 TWh by 2027 through bottom-up analysis of GPU shipments and power-per-chip data. Pilz et al. (2025) project global AI data center power demand could reach 327 GW by 2030, pointing out U.S. infrastructure bottlenecks that could shift AI development abroad. Roucy-Rochegonde and Buffard (2025) conclude global data center electricity consumption could reach 1100-2000 TWh by 2030, while Aljbour et al. (2024) forecast U.S. data center demand could hit 252–400 TWh by 2030. Chen (2025) highlights uncertainties in forecasting due to lack of corporate transparency. While these studies focus primarily on establishing power demand forecasts through extrapolation and expert insights, our work instead takes these forecasts as given and studies their implications for electricity prices and carbon emissions across multiple regions, including the United States, Europe, and China.

Research on AI's impact on electricity prices remains limited. Using elasticity estimates from the literature, Burian and Stalla-Bourdillon (2025) estimate that rising electricity demand from AI-driven data centers could increase gas prices by around 9% in Asia and Europe and 7% in the United States by 2026. Chandramowli et al. (2024) projects U.S. electricity demand could increase by 9% by 2028, potentially raising utility electricity costs by 19%. Unlike these studies, we employ a CGE model calibrated to projections of data center power demand through 2030. Our approach is also informed by observed trends in AI sectors. Specifically, it reflects how future ICT growth stems from both TFP improvements and increased capital investment, while capturing the rising electricity intensity within the ICT sector—patterns that mirror recent developments in AI sectors and companies.

Environmental impacts of AI have received increasing attention. Monserrate (2022) details cloud computing's ecological impacts in terms of water usage, electricity consumption, carbon emissions, noise pollution, and electronic waste. WEF (2025) estimates the electricity consumption breakdown across AI life cycle stages (development, training, deployment) and argues for energy scarcity as a key design principle for future AI infrastructure. In contrast, IEA (2025) highlights AI's growing data center energy demand alongside its promise for energy optimization and emissions reduction. Our work advances this literature by quantifying the specific carbon emission impacts of AI-driven data center expansion under various policy scenarios and infrastructure constraints using a CGE model.

Finally, a segment of the endogenous growth literature examines the potential role of

AI in driving economic development. Aghion et al. (2019) model AI as a novel automation technology that tackles tasks previously considered uniquely human, such as aspects of innovation itself. Their work suggests AI-driven growth may ultimately be constrained by essential tasks resistant to improvement—similar to Baumol's cost disease—potentially including energy and resource constraints. Epoch AI (2025) extend this thinking with a compute-centric endogenous growth model where AI progress, driven by computational power scaling, creates a feedback loop: enhanced AI capabilities boost output, which is partially reinvested into compute development, potentially generating annual growth rates up to 30%. While these theoretical approaches offer valuable insights into AI growth mechanisms, our multi-country multi-sector CGE model provides a more empirically grounded analysis by calibrating to historical data and IMF projections to assess the economic and environmental implications of AI expansion over the medium-term.

The remainder of the paper is structured as follows. Section 2 examines where AI production appears in the national accounts and analyzes these sectors growth patterns, their sources of growth, and their electricity usage. Section 3 introduces the IMF-ENV model and describes our scenario design, including assumptions about data center growth, alternative energy policies, and variations in medium-term constraints on renewable scale-up and transmission and distribution infrastructure. Section 4 presents our simulation results, analyzing the impacts on electricity prices, carbon emissions, and energy mix under different policy and infrastructure constraint scenarios. Section 5 concludes with a summary of key insights.

2 Growth, Productivity, and Electricity Use in AI-related sectors of the U.S. Economy

2.1 Sector Classification

AI production occurs primarily within NAICS sectors 518/519 (data processing, internet publishing, and other information services) and 5415 (computer systems design and related services), though it is not exclusive to these sectors. The AI production ecosystem comprises four distinct firm types: (i) specialized AI research labs (e.g., OpenAI, Anthropic); (ii) pure data center operators (e.g., Equinix); (iii) cloud services and infrastructure providers; and (iv) vertically integrated technology companies (e.g., Microsoft, Google) that span the entire value chain—from research through deployment to integration of AI with existing products like Google Search, Gmail, and Microsoft Office. These firms' core activities predominantly align with the aforementioned NAICS codes, making these sectors central to measuring AI-related economic activity.

To be precise, data centers are most often categorized as NAICS 518210 (Data Processing, Hosting, and Related Services) as they include, according to the US Census Bureau, activities such as application hosting, cloud storage services, computer data storage services, or computing platform infrastructure provision. For example, Equinix, one of the largest data center companies, is categorized under this industry code. Regarding the large and vertically integrated AI platform and service companies, as examples, META's NAICS code is 519290, while Alphabet operates under 519, 518 and 541511. Microsoft also has 518 and 541511 as one of its NAICS codes and IBM's codes are 5415 (54151 and 541512).

A few caveats are in order. First, our definition represents a narrower scope than broader classifications like the Information and Communications Technology (ICT) sector, which spans both manufacturing (computers, electronics) and services (telecommunications, software, IT services). It should be noted that the scenario simulations in IMF-ENV model are built around a TFP shock to the ICT sector, as the latter constitutes the smallest plausible proxy for the AI sector that can be lifted from the GTAPv11 database. The classification of AI here under NAICS codes 518/519 and 5415 also differs from, but overlaps with, the commonly used 'tech' category, which typically refers to several innovative technologies' companies with a very large market capitalization, ranging from hardware manufacturers (for example, Apple) to digital platform and service providers (Microsoft, Google, Meta, Alibaba) to essential component makers like semiconductor firms (Nvidia, TSMC, ASML). Hardware manufacturers and semiconductor firms are excluded here. Second, certain activities of AI companies are classified under traditional sectors, e.g., Equinix as a data center company also is a lessor of real estate (NAICS 531110), but such codes are excluded to avoid capturing non-AI activities. Third, AI production is becoming increasingly embedded across activities due to hybrid business models (e.g., Tesla investing in autonomous vehicles), among other reasons, making it difficult to make a clean one-to-one correspondence between AI-producing sectors and NAICS.

2.2 Data

Data on nominal and real value added (in SAAR 2017 USD for the latter) and real value added per employee (in SAAR 2017 USD per employee) were sourced from Haver Analytics, while data on the contributions of TFP and inputs (capital, labor, intermediates) to gross output growth were taken from the BEA-BLS Integrated Industry-level Production Accounts (KLEMS).

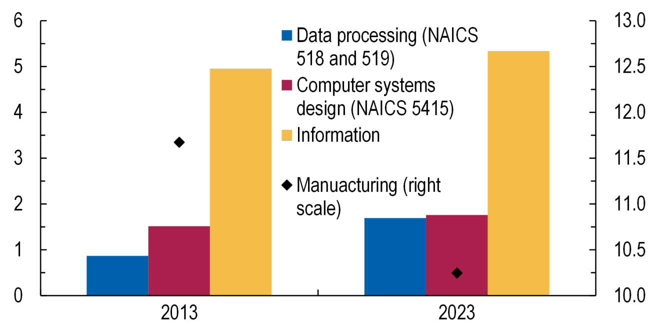
Limited transparency exists regarding electricity intensity in the AI production ecosystem due to minimal corporate disclosure (Chen 2025). To estimate electricity costs as a share of total expenditures, we gathered annual electricity consumption data from sustainability reports of publicly traded companies across three firm categories: AI platform and service providers, specialized data center operators, and semiconductor manufacturers. We calculated electricity costs by multiplying consumption figures by electricity prices from the US Energy Information Administration (EIA), using an average of industrial and commercial rates. This approach reflects that large tech companies likely benefit from lower industrial rates, while smaller data center operators typically face higher commercial rates as classified by the EIA. We obtained each company's total costs, that is, costs of revenues (or sales) plus operating expenses, from the annual 10-K reports they submitted to the Securities and Exchange Commission (SEC). We then calculated electricity's share of total costs by dividing electricity expenditures by the total costs. Finally, to calculate the average electricity share by company category, we take a weighted average based on companies' revenues also from their 10-K annual reports. It is worth highlighting, first, that two data center companies go private in the middle of the sample period and hence do not file 10-Ks or sustainability reports, and second, that for a very small number of companies electricity consumption is missing in their sustainability reports for part of our sample. To deal with missing data, we interpolate using the average annual growth rates for revenues and electricity consumption shares for each of the three company categories.

Finally, using 10-Q SEC filings, sourced from Capital IQ, we also compile data on capital expenditures by Microsoft, Alphabet, META, and Amazon between 2019 and 2024.

2.3 The Growing Macroeconomic Relevance of AI-producing Sectors

In the US, AI-producing sectors have experienced rapid growth, with their value-added increasing threefold from $372 billion (in constant 2017 USD) to $1.13 trillion between 2013 and 2023, thereby significantly outpacing both overall private industries and manufacturing. Consequently, these sectors' contribution to total nominal US GDP increased from 2.4 percent in 2013 to 3.5 percent in 2023, with the data processing sector nearly doubling its share during this period. In contrast, manufacturing's share of GDP declined by 1.5 percentage points over the same timeframe (Figure 1).

Figure 1: Share of AI-related Value-Added Output in GDP (percent of nominal GDP)



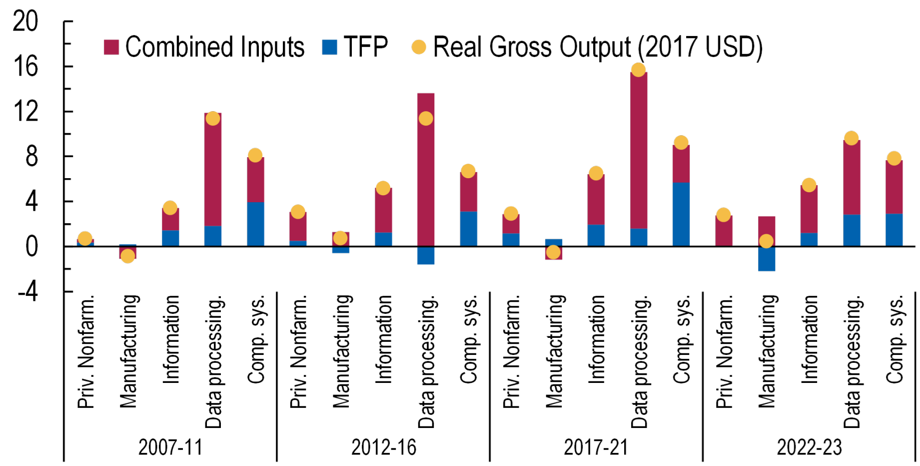
Sources: Haver Analytics. Note: NAICS=North American Industry Classification System.

From 2013 to 2023, AI-producing sectors increased their share of real US GDP from 2.1% to 5%, far outpacing their nominal GDP growth. This gap arises from opposing price trends: while the economy's overall price level, measured by the GDP deflator, climbed nearly 31% over the decade, prices in AI-related sectors like data processing and computer system design dropped by 17% and 27%, respectively. These declines reveal that, in real terms (adjusted for inflation), these sectors contribute more to the economy than their nominal figures suggest. Rapid productivity growth likely drives this, increasing supply faster than demand can keep up, pushing prices down as seen in agriculture, manufacturing, computer hardware, and solar photovoltaics—hinting that parts of AI production may already be commodifying.

Real gross output growth in the AI-producing sectors has outpaced the growth in private non-farm and manufacturing sectors (Figure 2). Moreover, the AI-producing sectors have demonstrated remarkable resilience during the global financial crisis of 07-09 and the 2020 pandemic-induced recession, maintaining positive and substantial growth rates while other industries stagnated or contracted. Growth in 2021-2022 has been particularly strong, with

AI-related services output expanding by 14.6%, far exceeding the overall private non-farm business sector growth of 5.1%. Similarly, the information sector, that contains data processing, recorded a growth rate of 7.7% growth in 2023 - the second highest across all industries after mining. Overall, the sectors involved in AI production have exhibited robust growth in value-added and gross output, increasing their importance in overall US output in the last decade.[1]

Figure 2: cost share in TFP, Combined Inputs, and Real Gross Output (percent)



Sources: Haver Analytics. Note: NAICS = North American Industry Classification System; TFP = Total factor productivity. Priv. Nonfarm= Private Nonfarm Business Sector.

Such rapid growth of AI-producing sectors has been driven by exceptional gains in labor productivity (LP). Specifically, value-added per employee in the information and AI-producing sectors grew nearly eight and thirteen times faster than overall economy-wide LP over the past decade, respectively. Within the information sector, the data processing sector demonstrated even larger LP gains. In contrast, manufacturing sector has experienced declines in average LP since 2007 (Figure 3). This differential growth has led to sectoral LP levels that far exceed the economy-wide average. The average LP in the data processing sector was $728 thousand (in constant and seasonally adjusted 2017 USD) in the first two quarters of 2024, approximately five times the national average, while for computer systems design it equaled $259 thousand, roughly twice as high as the overall average (Figure 4). The model in section 3

[1] The Bureau of Economic Analysis reached similar conclusions for a somewhat broader group of sectors it identifies as the "Digital Economy", which it studied under the Digital Economy Satellite Account (DASE) (Highfill and Surfield 2022). In addition to data processing and computer system design services, the DASE includes e-commerce, the digital components of manufacturing, and a smaller federal digital services sector.

Figure 3: Average Growth in Real Value Added per Employee (percent)

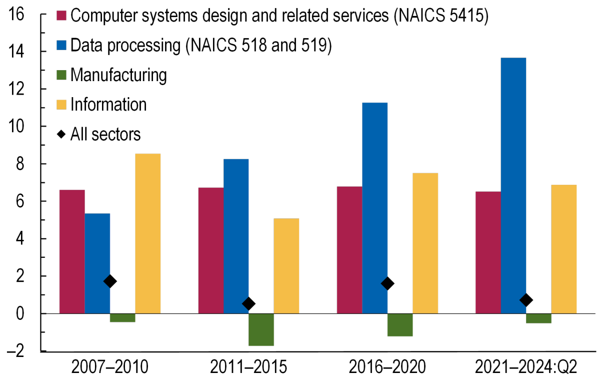
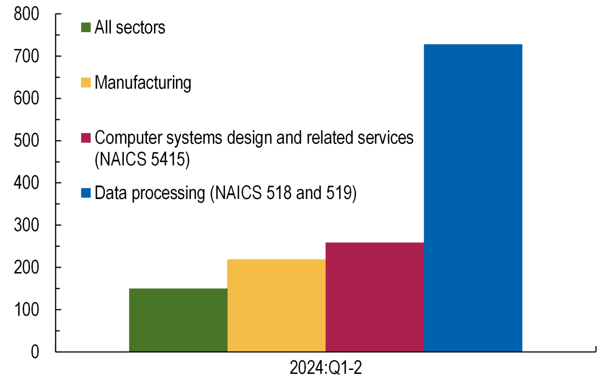


Figure 4: Real Value Added per Employee (SAAR thousand 2017 USD)



Sources: Haver Analytics. Note: NAICS=North American Industry Classification System.

builds upon the premise that rapid labor productivity gains in the information sector will likely continue.

The productivity growth in AI-producing sectors was largely the result of elevated investment in physical capital and the complementarity of intermediate inputs—contrary to computer systems design where labor and total factor productivity (TFP) contributed significantly to output growth (Figure 5). Hence, the remarkably high output-per-employee of data centers, relative to other sectors, is the result of fast capital accumulation which has required increased energy consumption as an intermediate input. Partly due to the increased capital-intensity of the data processing sector, the information sector showed net job losses in the first quarter of 2024 (Figure 6). The increased capital intensity is also apparent in the substantial capital expenditures made by major AI platforms and service providers (Figure 7). In the fourth quarter of 2024, Alphabet, Amazon, Meta, and Microsoft collectively spent nearly 75 billion USD on capital investments — a fivefold increase from the 15 billion USD recorded in the first quarter of 2019.

2.4 AI's demand for electricity

The analysis of major publicly traded companies in semiconductors, hardware, and software reveals significant variation in electricity costs as a share of total costs between 2019 and 2023 (Figure 8). Electricity costs make up 13–15 percent of total costs for data center companies, while they account for only 0.8–1.5 percent for semiconductor firms and AI service companies. However, this electricity intensity is rising rapidly among companies developing and

Figure 5: Contributions to Sectoral Gross Output Growth in AI-Producing Sectors (percent)

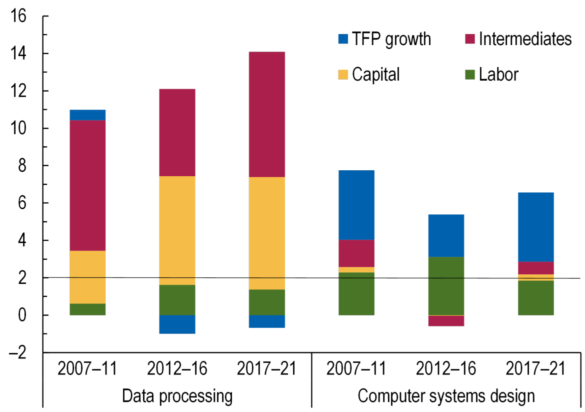
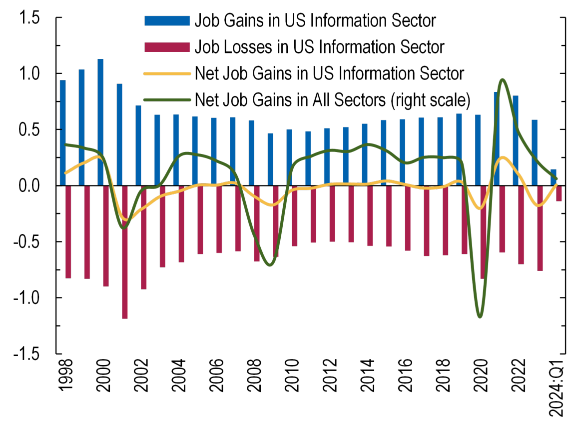


Figure 6: Job Gains and Losses in the U.S. Information Sector (millions)



Sources: Haver Analytics. Note: NAICS = North American Industry Classification System. BEA-BLS Integrated Industry-level Production Accounts (KLEMS). TFP = Total Factor Productivity. Priv. Nonfarm = Private Nonfarm Business Sector.

deploying AI models. On average, AI platforms and service companies have almost doubled their electricity cost share in less than five years (Figure 8). As these hyperscale companies increasingly integrate vertically by building, operating, and leasing their own data centers, their electricity cost shares will likely continue to grow. These observed and expected future empirical trends on electricity shares for AI platforms and service companies inform the calibration of the model presented in section 3.

The broader implications for global electricity consumption are substantial. Worldwide electricity consumption from data centers and AI is estimated to have reached 400-500 TWh in 2023, more than double the level in 2015, which had stayed mostly flat during 2015-2019 (OPEC 2024). For the U.S., where growth is the fastest, electricity demand from data centers is expected to increase from 178 TWh in 2024 to 606 TWh in 2030 under a medium-demand scenario (McKinsey 2024a). By 2030, AI-driven global electricity consumption could hit 1,500 TWh, conceivably making it comparable to India's current total electricity consumption, the third highest in the world. This projected electricity demand from AI by 2030 is around 1.5 times higher than expected demand from electric vehicles, another emerging source of electricity demand growth (Figure 9). While data centers currently account for about 1.5 percent of global electricity consumption, this share varies significantly by location. In the United States, data centers represent approximately 4 percent of total electricity use, with

Figure 7: Capital Expenditures by Selected AI Platform and Service Companies
(billion USD)

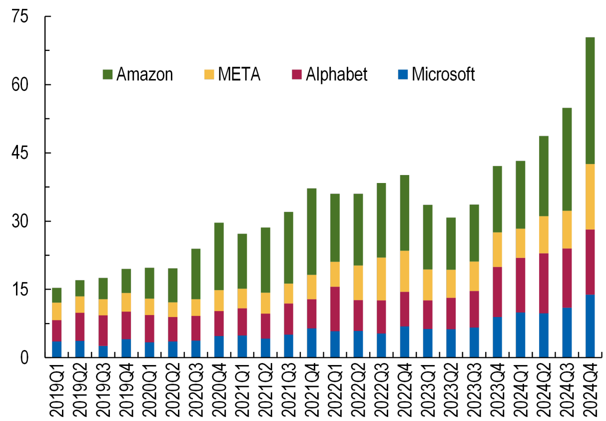
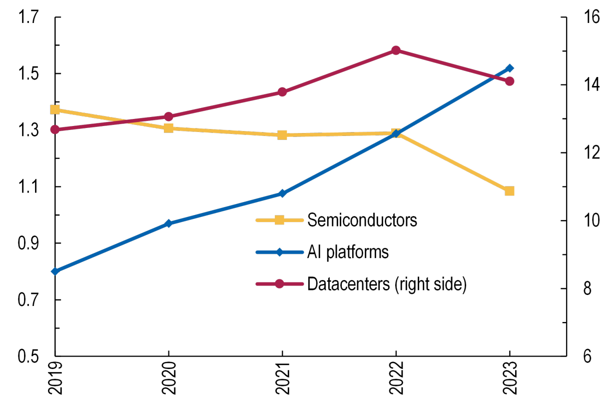


Figure 8: Estimated Electricity Costs for Publicly Traded Companies
(percent of total costs)



Sources: Figure 7: S and P Capital IQ. Figure 8: Companies' sustainability reports; and 10-K filings.

some regions showing particularly high concentration – notably Virginia, where data centers accounted for 26 percent of electricity consumption in 2024 (Electric Power Research Institute (EPRI) 2024; Shehabi et al. 2024).

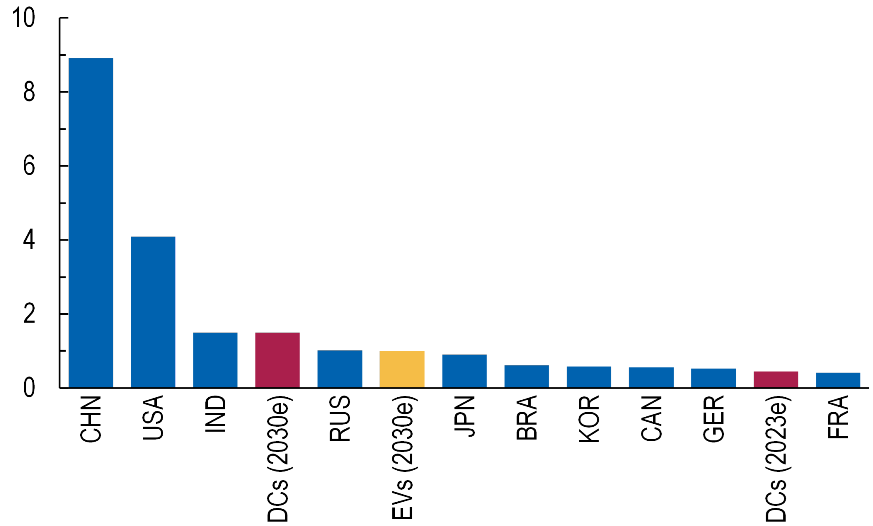
Recent developments in the AI industry have created more uncertainty around its future compute and energy demands. On the supply side, companies like DeepSeek are achieving breakthroughs in algorithmic efficiency that, combined with declining costs of ongoing hardware improvements, may lower the compute costs of AI models faster than previously anticipated. However, these efficiency gains may be counterbalanced by higher use of compute by companies pursuing better-performing models (Hoffmann et al. 2022). Adding to this complexity is the recent emergence of reasoning models – which require more compute in their deployment – and possibly greater AI use driven by lower costs and availability of open-source models.

3 AI and Energy Demand: An Application with IMF-ENV

3.1 Data center electricity demand forecasts

To assess the implications of rising electricity demand in AI-producing sectors, this exercise utilizes projected power consumption from data centers in three key regions—the United States, Europe, and China—between 2024 and 2030 (Figure 16). Aggregate level projections for

Figure 9: Electricity Demand for Data Centers Compared to Top Electricity Consuming Countries in 2023 (thousands of TWh)



Sources: EIA; and OPEC. Note: Estimates for data centers (DCs) and electric vehicles (EVs) are for the world and come from OPEC and IEA, respectively. Data labels in the figure use International Organization for Standardization (ISO) country codes.

these regions are derived from forecasts by McKinsey (2024a), McKinsey (2024b), and JP Morgan (2024). The projected annual growth rates in power demand are estimated at 22%, 13%, and 10% for the United States, Europe, and China, respectively. Specifically, the U.S. projection is based on McKinsey's "medium demand" scenario, while China's forecast is sourced from a JP Morgan study. For European countries, a GDP-weighted methodology was applied to the three largest economies—Germany, France, and Italy—which collectively account for approximately half of the region's total economic output. Additionally, the 2023 baseline power demand for China was assumed to be equivalent to that of the United States.

The forecasted US electricity consumption in 2030 used in the model's simulations is broadly in line with the US Department of Energy's (DOE) projected average consumption of 675 TWh when DOE's 2024-2028 growth rates are extended to 2030 (Shehabi et al. 2024). For China, the projected electricity consumption coming from data centers in 2030 stands on the lower end of the IEA's forecasted range of 260-470 TWh (IEA 2024). Finally, for the European Union and the UK, our projected 141 TWh of electricity consumption in 2030 is somewhat below the 205 TWh forecasted for the European Union in IEA (2024), when the 2022-2026 implied annual growth rate of 8% is extrapolated to 2030. Our projected annual growth rate of 13% is above theirs (at 8% between 2022 and 2026), but their starting point at 110 TWh in

2022 is above ours (60 TWh in 2023). However, McKinsey's data showed a consumption of 60 TWh for the EU and UK in 2023.

3.2 IMF-ENV: model basics

IMF-ENV is a multi-country dynamic Computable General Equilibrium (CGE) model developed at the IMF to analyze the intricate interactions among economic agents—households, firms, governments, and the external sector—across multiple sectors and markets. Its strength lies in capturing both direct and indirect effects of policy changes and economic shocks, making it a powerful tool for assessing general equilibrium outcomes at domestic and global levels. Another strength of the model is its inherent consistency: markets for all commodities and production factors must clear in each simulation period; all resource constraints are respected; and all macroeconomic balances (government budget, current account, and investment-savings equality) are maintained. This consistency is ensured through so-called "closure rules"—exogenous assumptions governing market clearing mechanisms—which also link these balances to external projections from the World Economic Outlook. As such, IMF-ENV provides a robust framework for medium- and long-term policy analysis. It is particularly well-suited for evaluating structural shifts in the economy that could arise from energy policies, climate policies and trade reforms.

Built on neoclassical optimization principles and competitive market assumptions, in this analysis IMF-ENV simulates the global real economy with a recursive dynamic structure extending to 2030. Agents' responses regarding consumption, production, and trade are driven by different elasticities. There are four factors of production: labor, capital, land, and natural resources, with capital distinguished by a vintage structure (i.e., old versus new). Using the GTAPv11 Power database (Aguiar et al. 2022; Chepeliev 2023), the model is calibrated for 25 regions, including the G20 countries, and 36 sectors. Energy is a key focus of the model, divided into electric (e.g., solar, wind, nuclear, hydropower, coal, oil, gas and rest) and non-electric (e.g., coal, oil, gas extraction) sectors, with GHG emissions tied to direct fossil fuel consumption for any economic activity. Model regions are connected to one another through bilateral trade flows that are modeled with the Armington specification (Armington 1969), where demand for goods is differentiated by region of origin. This trade specification takes into account bilateral trade flows while considering differences in prices, transportation, and

Figure 10: CES production function of the AI sector

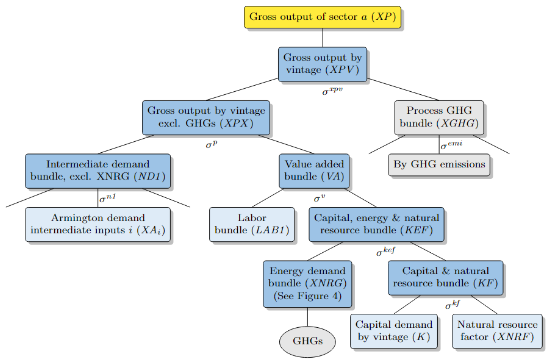
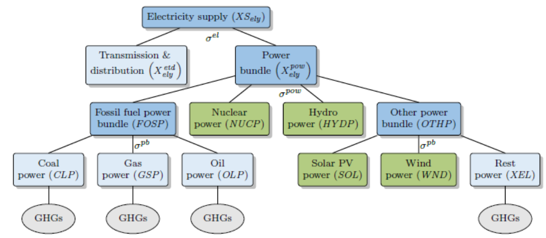


Figure 11: CES production function of the Power sector



Sources: IMF-ENV model. Further details on the model are available in Chateau et al. 2025. In the right panel, electricity generation technologies marked in green denote technologies that do not emit GHGs.

trade costs by commodity by trading partners.The trade structure allows IMF-ENV to model complex interdependencies within economies, and assess how structural shifts in one region can transmit to rest of the world through bilateral trade networks.

In IMF-ENV, separate production functions are defined for each economic sector or activity. Sector-specific representative firms minimize their production costs under the assumption of constant returns to scale, implying that each sector operates in perfectly competitive markets. The production function in each sector consists of nested constant-elasticity-of-substitution (CES) functions, which capture various substitution possibilities between different pairs of input bundles. The nested CES system represents the optimization process where each representative firm minimizes the cost of purchasing intermediate inputs and production factors within the constraints of the production function. Figure 10 shows the CES nesting of non-agricultural activities, including the AI sector.

In IMF-ENV, the standard configuration dictates that each economic activity produces a single commodity, with the exception being the electricity generation sector. A notable characteristic of IMF-ENV is its differentiation among eight distinct electricity generation technologies: coal, gas, oil, nuclear, hydro, solar PV, wind, and others (including geothermal, biofuels, tidal, and waste technologies). Consequently, electricity generation activities follow a many-to-one mapping, where all power generation activities are used to produce a single electricity commodity as shown in Figure 11. Importantly, the intermediate inputs from the transmission and distribution (T&D) sector are necessary for scaling up power generation from any source.

Therefore, any expansion of power generation must be accompanied by corresponding growth in the T&D sector. [2] Further details on the different structural and behavioral assumptions of the model are available in (Chateau et al. 2025).

3.2.1 Static and dynamic calibration

The first step in the calibration process entails calibrating the model to the 2017 base year data from the GTAPv11 database. To this end, values for key parameters, such as elasticities of trade, consumption (income), and production, are sourced from the literature and the GTAPv11 database. Next, the CES factor share parameters of all the production functions are then calculated so that the model replicates the 2017 base-year data. To simulate the baseline scenario, several parameters must be calculated during the dynamic calibration process with the goal of projecting several exogenous drivers. Here we describe the key steps. First, demographic trends and labor force participation rates are taken from the WEO database to project labor supply. Second, the labor productivity path for each country is then calibrated in an iterative process to match real GDP growth rates from the IMF's WEO projections. Third, the share of each type of electricity technology is controlled by dynamic calibration of the CES share parameters using projections from (Keramidas et al. 2025). Fourth, CO_2 emissions are calibrated by an emissions shifter also based on (Keramidas et al. 2025). Finally, various closure rules maintain macroeconomic balances: (i)-(ii) the government budget balance and the current account balance (CAB), both as a share of GDP, are assumed to follow the WEO projections; (iii) investments are driven by the sum of consumer savings (as a share of GDP), government savings (which follow exogenous projections), and foreign savings (linked to the CAB closure rule). This calibration procedure enables the model to replicate historical data while projecting plausible future paths under varying conditions.

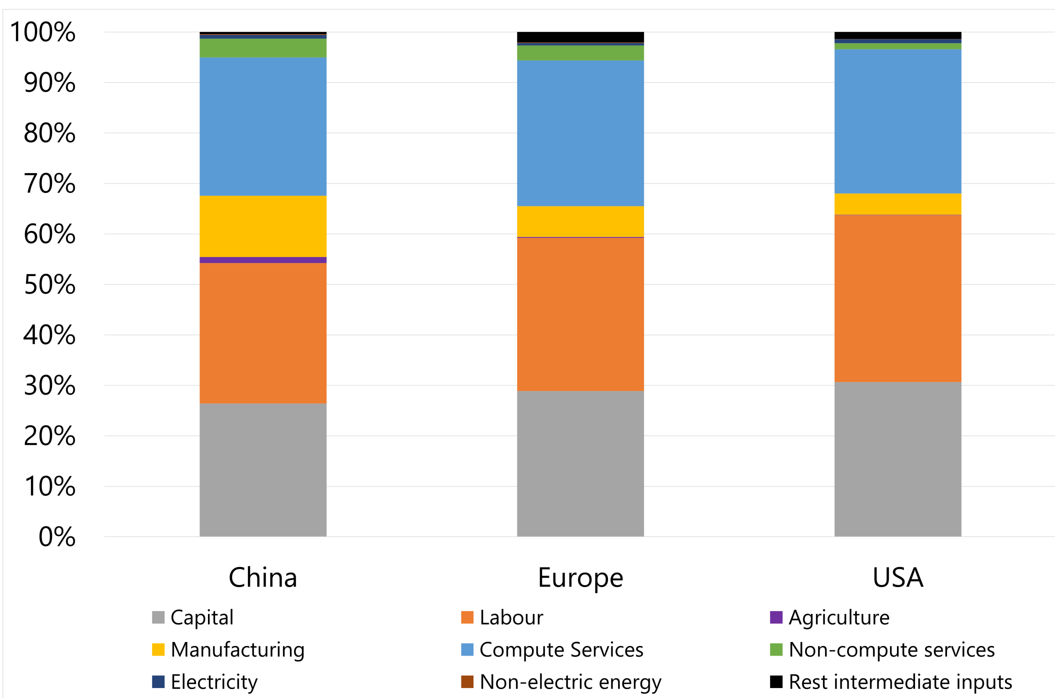
3.3 Scenarios

In this paper we simulated three scenarios with IMF-ENV. The *baseline* scenario does not account for AI growth trends, and therefore the energy and emissions trends are calibrated

[2] This assumption implies that all new generation must be connected to the grid. Consequently, the model does not include off-grid capacity additions, which continue to be a minor share of total generation in all countries where AI-related shocks are simulated.

solely based on policies that were implemented until the year 2024. Based on data from the GTAPv11 database, the global average input cost shares for labor and capital in the Information Technology sector are approximately 30 percent. About a quarter of intermediate inputs are from other compute services, and roughly 5 percent are from manufacturing. Energy, which mainly consist of electricity, is our key input of interest and constituted about 1 percent of the input costs in 2017. These cost shares are shown in Figure 12 and are broadly similar in the U.S., China and Europe. Recent data shows that in less than five years, AI platforms and service companies have increased their electricity cost share from 0.8 percent in 2019 to 1.5 percent in 2023 (Figure 8). In our simulations, we assume that the increasing trend in IT sector's electricity intensity will persist in the United States, which is expected to experience the most significant AI expansion. Under this assumption by 2030, this intensity rises to 4 percent, up from 1 percent in 2017 in the U.S. For all the other countries these shares are kept identical to the 2017 values.

Figure 12: Cost shares of inputs in the AI sector (percent, 2017)



Sources: IMF-ENV based on Aguiar et al. 2022

In addition to the baseline, we model two AI growth scenarios. In both AI scenarios, we introduce an AI-driven total factor productivity (TFP) shock within the IT sector. This shock is calibrated to ensure that the sectoral electricity demand from the IT sector aligns with the forecasts presented in Table 1. The TFP shock is applied to the Value Added (VA) bundle within the nested-CES production function (see Figure 10) and is specifically directed at production processes that incorporate new capital. The AI impact remains unchanged in both scenarios of AI development; however, the variation is due to electricity sector policies, which influence the composition of electricity supply. The first AI scenario, *AI under current energy policies*, presents an AI shock under current energy policies that are consistent with the baseline, assuming no changes in the electricity generation mix. Differently, in the second AI scenario, *AI under alternative energy policies*, additional supply-side measures are used to increase renewables' share through feed-in tariffs aligned with regional long-term strategies following NDC-LTS (Keramidas et al. 2025). Understandably, the advancement of AI technology is heavily reliant on the growth of electricity supply. Therefore, energy policies should prioritize stimulating the supply side. Among the supply-side policies, we implement feed-in tariffs for renewable energy in the second AI scenario because this incentive has been historically adopted by all our target countries within their policy frameworks, and renewable technologies represent some of the most cost-effective options available.[3]

Several factors could potentially slow the growth of solar and wind capacity in the U.S. over the next five years, including supply policy uncertainty, chain constraints, delays in permitting processes, and fluctuations in commodity prices. Additionally, these factors could also impact new investments in updating and expanding the grid, which may contribute to limited expansion of renewable energy. For both AI scenarios, simulations are also done with different assumptions concerning medium-term constraints that could limit the growth of AI sectors. The sensitivity of the model results is checked against the following assumptions - (1) Growth potential of renewables between 2025-2030 (*Current/Alternative policies with smaller renewables scale-up*), and (2) Investments in transmission and distribution infrastructure (*Current/Alternative policies with no additional investments in T&D*).

In the IMF-ENV model, the first assumption is addressed by introducing a constraint that caps the increase in sectoral production levels of solar PV and wind power generation such that the annual growth rates are equal to or below the average growth rates seen in the last five years. In IMF-ENV, all power generation expansion needs to be supported by com-

[3] Alternate policy instruments like feebates and regulations can also be used and calibrated to deliver similar results in the power sector, however, with different macroeconomic, price and emission impacts (Chateau et al. 2024).

plementary expansion of transmission and distribution (T&D) sector. The second constraint is modeled by adding a new constraint that fixes the sectoral investment pathway of the T&D sector to that in the baseline pathway. It is important to note that investments in T&D sectors are increasing in the baseline scenario and therefore, this additional constraint highlights a situation where the sector's expansion does not sufficiently keep up with the increase in new power generation capacity that is added in the economy. In all scenarios presented in the paper, power generation from hydropower and nuclear technologies is capped at baseline generation levels as expansion of these technologies largely depends on political decisions and geographical capacity rather than market mechanisms.

4 Results

The TFP shock in the AI sector improves sectoral productivity, resulting in higher output from this sector. This leads to increased demand for all inputs, including electricity, and the AI shock increases electricity consumption by the IT sector in the U.S., Europe, and China. The increased demand for electricity can be addressed through two primary methods: (1) increasing the overall electricity production within the economy, and (2) reallocating electricity resources from other economic activities to the AI sector. In the former channel, power producers expand total generation, which would come from both carbon-intensive sources like natural gas and coal, and zero-carbon sources like solar, wind and other renewables. However, the composition of electricity generation by technologies varies across countries and is based on their relative production costs and current policies. The latter channel is influenced by the extent to which electricity expansion can be achieved and the relative productivity differences across various sectors within a region.

By 2030, in the *AI scenario under current energy policies*, total electricity supply increases by 8 (525 TWh), 3 (145 TWh) and 2 (237 TWh) percent in the U.S., Europe, and China, respectively, relative to the baseline scenario. In the *AI scenario under alternative energy policies*, the increase in total electricity supply is kept the same, but its composition shifts in favor of renewables. In the U.S., Europe, and China, generation from solar and wind sources offsets about 58, 35 and 166 TWh of generation from other sources, including largely natural gas in the U.S and coal power in China (Figure 13).

Figure 13: Electricity Supply and Generation Mix, 2030 (TWh)

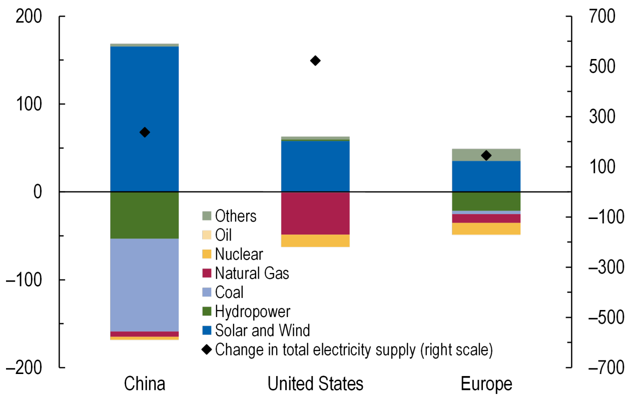
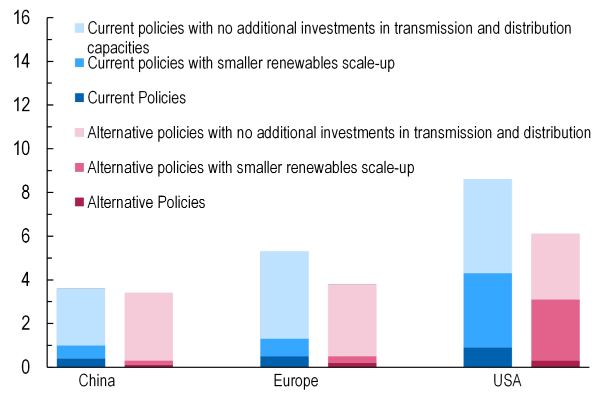


Figure 14: Change in Electricity Prices Relative to the Baseline Scenario, 2030 (percent)



Sources: Figure 4: IMF-ENV model. Note: The left axis shows the change in generation mix under alternative energy policies relative to current policies in terawatt hours (TWh). Feed-in tariffs increase generation from solar and wind sources. The right axis shows the total increase in electricity supply relative to the baseline scenario in TWh, which is identical in both current energy policies and alternative energy policies. Figure 5: IMF-ENV model.

In both scenarios, due to rising marginal costs of electricity supply, the increase in generation is less than proportional to economy-wide demand growth, driving electricity prices up. In this case, the surge would be 0.9, 0.45 and 0.35 percent in the U.S., Europe, and China respectively, under current energy policies (Figure 14). The electricity price increase is less significant under *alternative policies* because of the feed-in tariff on solar and wind. This tariff lowers the generation cost of these technologies, which to begin with have relatively low production costs and a higher share in total electricity generation compared to current policies.

The impact on electricity prices is particularly sensitive to medium-term limitations that may impede the expansion of power generation capacity. Furthermore, price pressures may also originate from other factors such as the increased electrification of economic activities and the adoption of electric vehicles, which are not modeled in this paper. Material pressure on the prices would be added should the expansion of renewable energy sources decelerate compared to recent historical trends, and additional investments in transmission and distribution infrastructure be absent, relative to the baseline. With the realization of these two constraints, the AI expansion could alone lead to a price increase in the U.S., Europe, and China under current policies, potentially escalating to 8.6, 3.6, and 5.3 percent, respectively, by 2030 (Figure 14).

Among these two constraints, modeling results indicate that the capacity of the grid infrastructure is a more critical factor contributing to price increases.[4] Without further investments in transmission and distribution, supporting the expansion of the AI sector would require redirecting electricity from other economic activities. This shift would pose significant challenges especially for energy-intensive manufacturing sectors. For example, in the U.S. the annual growth in these sectors' value added would experience an average reduction of 0.3 percent point compared to the baseline scenario, negatively impacting annual GDP growth by 0.1 percent point.

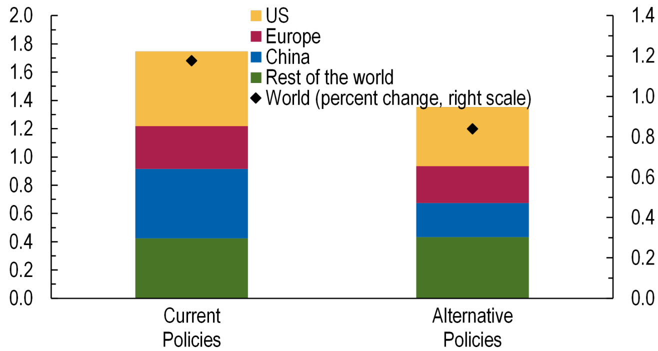
In both AI scenarios, global and regional GHG emissions increase due to the increased energy demand resulting from the expanded IT sector and its spillovers to the economy. Under current energy policies, the 2030 increase is 5.5, 3.7 and 1.2 percent in the US, Europe and China, respectively, with a global average increase of 1.2 percent (Figure 15). In cumulative terms, this translates into a global GHG emissions increase of 1.7 Gt between 2025 and 2030, which is similar to Italy's energy-related GHG emissions over a 5-year period. Notably, under alternative energy policies, even a modest decarbonization of the power sector limits the total cumulative global GHG emission increase from the AI shock significantly to 1.3 Gt by 2030, 24 percent less global emissions than the under current energy policies.[5]

Lastly, the TFP shock on the AI sector has a positive impact on GDP levels. Under *current energy policies*, the AI shock raises the average annual growth rate of global GDP by 0.5 percentage point between 2025 and 2030. Given our choice of calibrating the AI shock, the GDP gains are more significant in countries where the projected growth rate of the IT sector and the sectors' relative importance in the economy are higher. Under alternative energy policies, these benefits are slightly diminished due to the implementation of the feed-in tariff policy. The fiscal impact of these tariffs ranges from 0.3 to 0.6 percent of GDP across various countries. In the simulations, it is assumed that this cost is financed through increased lump-sum taxes, resulting in a slight reduction in household consumption. The growth in GDP level from AI expansion greatly exceed the fiscal costs, resulting in similar gains in annual GDP

[4]Given the strong commitment of major AI players, many medium-term power supply rigidities could be overcome, leading to a small increase in electricity prices. Public investments are being made in the United States for upgrading transmission and distribution infrastructure to meet rising electricity demand. Innovative solutions like power coupling, see Engel et al. (2025), and small modular nuclear reactors could offer flexibility, making constraints less restrictive than expected. Most new nuclear capacity is expected online no earlier than the early 2030s.

[5]This estimate is conservative compared with that of Stern and Romani (2025) who project that AI's energy demand could contribute between 0.4 and 1.6 Gt of carbon dioxide equivalent annually by 2035.

Figure 15: Emission Impacts of Expansion in IT Sector (Cumulative greenhouse gas emissions, $MtCO_2e$; percent change, relative to baseline)



Sources: IMF-ENV model. Note: The left axis shows the total greenhouse gas emissions increase in metric tons of carbon dioxide equivalent ($MtCO_2e$) between 2025 and 2030 resulting from information technology (IT) sector expansion in selected regions. The right axis shows the total increase in global emissions in 2030 relative to the baseline emissions as a result of this expansion.

growth both AI scenarios we modelled.

In sum, while the AI-induced expansion of the IT sector is expected to raise global GDP, the development also comes at a cost in terms of higher carbon emissions. Drawing on a median social cost of carbon (SCC) estimate of $39 per ton—based on 147 published studies with over 1,800 estimates (see Moore et al. (2024))—the additional social cost of 1.3 to 1.7 Gt of carbon-equivalent emissions is about $50.7 to $66.3 billion, or 1.3 to 1.7 percent of the AI-driven increase in real world GDP between 2025-2030.

5 Conclusions

As AI technologies continue to evolve and proliferate, the demand for computational power and electricity is poised for a significant surge. AI-related electricity consumption could reach up to 1,500 TWh by 2030, possibly outpacing other emerging sources of demand, like electric vehicles, and becoming comparable to India's total electricity consumption, which is the third largest in the world.

The increasing electricity demand from the IT sector will stimulate overall supply, which—if sufficiently responsive—will lead to a small increase in electricity prices. More sluggish supply responses, especially in expanding medium-term renewables capacity and transmission and distribution infrastructure, will lead to much stronger price surges, impacting households and businesses, and possibly constraining the growth of the AI industry itself. In the U.S., the country expected to experience the largest surge in AI-driven demand for electricity, the AI expansion alone could increase electricity prices by up to 9 percent, adding to price pressures coming from many other sources.

In addition, under current energy policies, the AI-driven rise in electricity demand could add 1.7 Gt in global greenhouse gas emissions between 2025 and 2030, similar to Italy's energy-related GHG emissions over a 5-year period. The social cost of these additional emissions represents only a very small portion of the anticipated aggregate economic benefits from AI. However, they would nonetheless contribute to an already concerning accumulation of emissions. In addition, while the additional emissions will have global impacts, the benefits of AI will likely be unequal both across countries and among different groups within societies, potentially exacerbating existing inequalities (Cazzaniga et al. 2024, Cerutti et al. 2025).

Demand for compute and electricity from AI service producers is subject to wide uncertainty. The emergence of more efficient, open-source AI models, such as DeepSeek, has added to the uncertainty, as algorithmic improvements tend to reduce compute costs and electricity demand. At the same time, lower compute costs stimulate AI use, which—together with the development of more energy-intensive reasoning models—adds upward electricity demand pressure. This heightened uncertainty poses a risk of delaying crucial energy investments, potentially resulting in underinvestment and escalating energy prices.

Implementing policies that incentivize renewables can enhance electricity supply, help mitigate price surges and reduce the emission impacts. Ultimately, such measures will enable the realization of AI's full potential in a sustainable manner.

References

Aghion, Philippe, Benjamin F. Jones, and Charles I. Jones (2019). "9. Artificial Intelligence and Economic Growth". *The Economics of Artificial Intelligence*. Ed. by Ajay Agrawal, Joshua Gans, and Avi Goldfarb. Chicago: University of Chicago Press, pp. 237–290. ISBN: 9780226613475.

Aguiar, A., M. Chepeliev, E. Corong, and D. van der Mensbrugghe (2022). "The GTAP Data Base: Version 11". *Journal of Global Economic Analysis* 7.2, pp. 1–37.

Aljbour, Jordan, Tom Wilson, and P Patel (2024). "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption". *EPRI White Paper no. 3002028905*.

Armington, P. S. (1969). "A Theory of Demand for Products Distinguished by Place of Production". *International Monetary Fund Staff Papers* 16, pp. 159–76.

Burian, Vlad and Arthur Stalla-Bourdillon (2025). "The increasing energy demand of artificial intelligence and its impact on commodity prices". *Economic Bulletin Boxes* 2.

Cazzaniga, Mauro, Florence Jaumotte, Longji Li, Giovanni Melina, Augustus J Panton, Carlo Pizzinelli, Emma J Rockall, and Marina Mendes Tavares (2024). *Gen-AI: Artificial intelligence and the future of work*. International Monetary Fund.

Cerutti, Eugenio, Antonio Garcia Pascual, Yosuke Kido, Longji Li, Giovanni Melina, Marina Mendes Tavares, and Philippe Wingender (2025). *The Global Impact of AI: Mind the Gap*. IMF Working Paper 25/XX. International Monetary Fund.

Chandramowli, Shankar, Patty Cook, Justin Mackovyak, Himali Parmar, and Maria Scheller (2024). *Power Surge: Navigating US Electricity Demand Growth*. New York: ICF.

Chateau, Jean, Florence Jaumotte, and Gregor Schwerhoff (2024). "Climate Policy Options: A Comparison of Economic Performance". *Energy Policy* 192.

Chateau, Jean, Hugo Rojas-Romagosa, Sneha Thube, and Dominique van der Mensbrugghe (2025). *IMF-ENV: Integrating Climate, Energy, and Trade Policies in a General Equilibrium Framework*. IMF Working Paper 25/77. International Monetary Fund.

Chen, Sophia (2025). "How much energy will AI really consume? The good, the bad and the unknown". *Nature* 639, pp. 22–24.

Chepeliev, Maksym (2023). "GTAP-Power Data Base: Version 11". *Journal of Global Economic Analysis* 8.2.

Comunale, Mariarosaria and Andrea Manera (2024). "The economic impacts and the regulation of AI: A review of the academic literature and policy actions".

Cushman and Wakefield (2024). *Global Data Center Market Comparison 2024*.

De Vries, Alex (2023). "The growing energy footprint of artificial intelligence". *Joule* 7.10, pp. 2191–2194.

Electric Power Research Institute (EPRI) (2024). *Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption*.

Engel, Alex, David Posner, and Uday Varadarajan (2025). *How "Power Couples" Can Help the United States Win the Global AI Race*. Report. RMI.

Epoch AI (2025). *GATE: Modeling the Trajectory of AI and Automation*. Accessed: April 8, 2025.

Farrell, Henry and Abraham Newman (2023). *Underground empire: How America weaponized the world economy*. Random House.

Highfill, Tina and Christopher Surfield (2022). *New and Revised Statistics of the U.S. Digital Economy, 2005–2021*. Report. Bureau of Economic Analysis.

Hoffmann, Jordan, Sebastien Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, et al. (2022). "Training Compute-Optimal Large Language Models". *arXiv preprint arXiv:2203.15556*.

International Energy Agency (2024). *World Energy Outlook 2024*.

— (2025). *Energy and AI*. Tech. rep. Licence: CC BY 4.0. Paris: IEA.

JP Morgan (2024). *How AI is Shaping These Three Industries in China*.

Keramidas, Kimon, F. Fosse, F.J. Aycart Lazo, P. Dowling, R. Garaffa, J. Ordonez, Petrovic, et al. (2025). *Global Energy and Climate Outlook 2024*. JRC139986. Luxembourg: Publications Office of the European Union.

Korinek, Anton and Jai Vipra (2024). *Concentrating Intelligence: Scaling and Market Structure in Artificial Intelligence*. Tech. rep. 33139. Cambridge, MA: National Bureau of Economic Research.

McKinsey (2024a). *How Data Centers and the Energy Sector Can Sate AI's Hunger for Power*. Article, September 17.

— (2024b). *The Role of Power in Unlocking the European AI Revolution*. Article, October 24.

Monserrate, Steven Gonzalez (2022). "The Cloud Is Material: On the Environmental Impacts of Computation and Data Storage". *MIT Case Studies in Social and Ethical Responsibilities of Computing* Winter 2022. https://mit-serc.pubpub.org/pub/the-cloud-is-material.

Moore, Frances C., Moritz A. Drupp, James Rising, Simon Dietz, Ivan Rudik, and Gernot Wagner (2024). *Synthesis of Evidence Yields High Social Cost of Carbon due to Structural Model Variation and Uncertainties*. Tech. rep. 32544. Cambridge, MA: National Bureau of Economic Research.

OPEC (2024). *World Oil Outlook 2050*.

Pilz, Konstantin F, Yusuf Mahmood, and Lennart Heim (2025). “AI’s Power Requirements Under Exponential Growth: Extrapolating AI Data Center Power Demand and Assessing Its Potential Impact on US Competitiveness”. *RAND Corporation: Santa Monica, CA, USA*.

Roucy-Rochegonde, Laure de and Adrien Buffard (2025). *AI, Data Centers and Energy Demand: Reassessing and Exploring the Trends*. Ifri Papers. Ifri.

Shehabi, Arman, Sarah J. Smith, Alex Hubbard, Alex Newkirk, Nuoa Lei, Md Abu Bakar Siddik, Billie Holecek, et al. (2024). *2024 United States Data Center Energy Usage Report*. Tech. rep. LBNL-2001637. lbnl-2024-united-states-data-center-energy-usage-report.pdf. Berkeley, California: Lawrence Berkeley National Laboratory.

Stern, Lord Nicholas and Mattia Romani (2025). *What is AI’s role in the climate transition and how can it drive growth?* World Economic Forum.

World Economic Forum (Jan. 2025). *Artificial Intelligence’s Energy Paradox: Balancing Challenges and Opportunities*. https://reports.weforum.org/docs/WEF_Artificial_Intelligences_Energy_Paradox_2025.pdf. Accessed April 4, 2025.

Appendix A: Tables and figures

Figure 16: Projected Power Demand From AI and Data Centers

Metric		Region	2023	2024	2025	2026	2027	2028	2029	2030	CGAR	Sources
Electricity consumption from data centers	(TWh)	US	147	178	224	292	371	450	513	606	22%	McKinsey
		EU+UK	60	68	77	87	98	111	125	141	13%	McKinsey
		Germany	12	14	16	18	20	23	26	29	13%	McKinsey & IMF staff calculations
		France	8	9	11	12	14	15	17	20	13%	McKinsey & IMF staff calculations
		Italy	6	7	8	9	10	12	13	15	13%	McKinsey & IMF staff calculations
		ROEU	24	27	30	34	39	44	49	56	13%	McKinsey & IMF staff calculations
		China	147	162	178	196	215	237	260	286	10%	JPMorgan & IMF staff calculations
	(GW)	US	19	23	28	37	47	57	65	77	22%	McKinsey
		EU+UK	8	9	10	11	12	14	16	18	13%	McKinsey
		Germany	2	2	2	2	3	3	3	4	13%	McKinsey & IMF staff calculations
		France	1	1	1	2	2	2	2	2	13%	McKinsey & IMF staff calculations
		Italy	1	1	1	1	1	1	2	2	13%	McKinsey & IMF staff calculations
		ROEU	3	3	4	4	5	6	6	7	13%	McKinsey & IMF staff calculations
		China	19	21	23	25	27	30	33	36	10%	JPMorgan & IMF staff calculations
		World	51-63			127				190		OPEC WOO 2024

Note: Blue figures are forecasts or estimates.

PUBLICATIONS

Power Hungry: How AI Will Drive Energy Demand

Working Paper No. WP/2025/081