COMMODITY SPECIAL FEATURE: ONLINE ANNEX 1.1

Power Hungry: How AI Will Drive Energy Demand

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

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 datacenter 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, datacenters 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 datacenter 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 datacenter 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.

Data

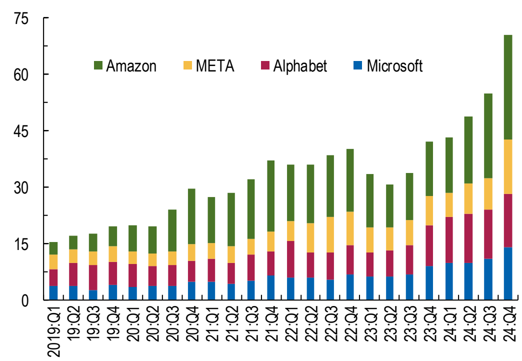
Data on real value added (in SAAR 2017 USD) and real value added per employee (in SAAR 2017 USD per employee) were sourced from Haver Analytics, while data on the contributions of TFP and

Version date: April 20, 2025, 6:21 p.m.

inputs (capital, labor, intermediates) to gross output growth were taken from the BEA-BLS Integrated Industry-level Production Accounts (KLEMS).

For Online Annex Figure 1.1.3, we compile yearly electricity consumption for publicly traded firms from their publicly available sustainability reports. Using an average between the industrial and commercial yearly retail prices of electricity from the US Energy Information Administration (EIA), we calculate estimated total costs for the companies' electricity consumption. We use an average between industrial and commercial prices of electricity because large technological companies probably face the lower industrial prices of electricity, while other smaller datacenter companies are categorized as commercial by the EIA and probably face a higher price commercial price of electricity. Next, we collect total costs (costs of revenues or sales plus operating expenses) for each company from their 10-K annual reports filed with the US Securities and Exchange Commission (SEC). Having costs from electricity consumption and total costs allows us to calculate the former as a share of the latter. 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 datacenter 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.

Online Annex Figure 1.1.1. Capital Expenditures by Selected AI Platform and Service Companies
(Billion USD)



Sources: S&P Capital IQ; and IMF staff calculations.

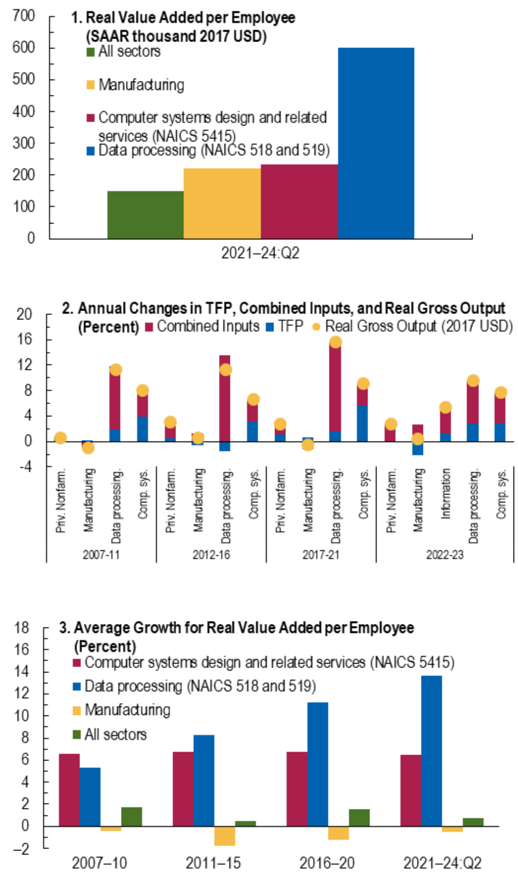
Using 10-Q SEC filings, we also compile data on capital expenditures by Microsoft, Alphabet, META, and Amazon between 2019 and 2024 (Figure 1). The total capital expenditures by these four large companies have quintupled in the last five years and has more than doubled from 2023Q3.

Descriptive analysis

The main text of the commodity special feature documents the growing macroeconomic importance of the AI-related sectors, with their share of nominal GDP growing from 2.3% to 3.5% between 2013 to 2023. This growth of AI's economic footprint is rooted both in rapid labor productivity growth and labor absorption, as documented in Figure 2, panel 2 and Figure 1. SF.2 (panel 2) of the main text respectively. While not reported in the main text, this growth has resulted in sectoral labor productivity (LP) levels that substantially exceed those of the average sector in the U.S. economy: data processing LP is now approximately four times the national average, while computer systems design LP is about 50 percent higher (Figure 2, panel 1).

These sectors also demonstrated notable resilience during the 2020 pandemic recession and the 2008-2009 crisis, with positive and substantial value-added growth rates (Figure 2, panel 2). Post-pandemic recovery has been particularly strong, with AI-related services output expanding by 11.7%, far exceeding the overall business sector growth of 4%. Within this trend, 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.

Online Annex Figure 1.1.2. The Growing Macroeconomic Relevance of AI-Producing Sectors



Sources: Haver Analytics; and IMF staff calculations.
Note: NAICS = North American Industry Classification System; TFP = total factor productivity. Priv. Nonfarm= Private Nonfarm Business Sector.

Overall, the sectors involved in AI production, data processing and computer system design, have exhibited robust growth in value-added and gross output, increasing their importance in overall US output in the last decade.

Part II. AI and Energy Demand: An Application with IMF-ENV

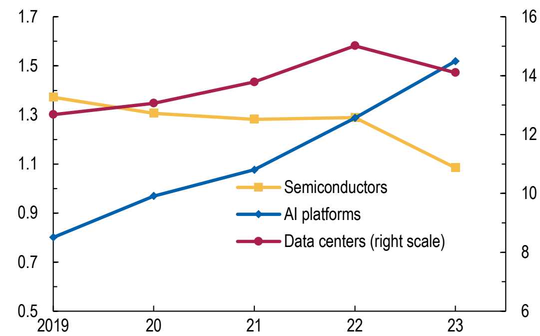
Datacenter electricity demand forecasts

To assess the implications of rising electricity demand in AI-producing sectors, this exercise uses projected power consumption from data centers in three key regions—the United States, Europe, and China—between 2024 and 2030 (Online Table 1.1.1). Aggregate level projections for these regions are derived from forecasts by McKinsey and JP Morgan. 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 datacenters in 2030 stands on the lower end of the IEA's forecasted range of 260-470TWh (IEA, 2025). Finally, for the European Union and the UK, similarly to us, IEA (2024) projects roughly a doubling of the electricity consumption coming from datacenters in the EU between 2022 and 2030, when the 2022-2026 growth rate is extended to 2030.

Online Annex Figure 1.1.3. Estimated Electricity Costs for Publicly Traded Companies
(Percent of total costs)



Sources: Companies' sustainability reports; 10-K filings; and IMF staff calculations.

Online Annex Table 1.1.1. 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.

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. The strength of the model 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 database (Aguiar et al. (2022)), the model is calibrated for 25 regions 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 fuel consumption for any economic activity. The 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. Further details on the model are available in Chateau et al. (forthcoming).

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 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 and others (2024). Fourth, CO_2 emissions are calibrated by an emissions shifter also based on Keramidas and others (2024). 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.

Scenarios

Three core scenarios are simulated for this exercise. The *baseline* scenario excludes AI influences, projecting energy and emissions trends based on current policies that were enacted up to 2024. In the *baseline*, the regional electricity mix in IMF-ENV is determined based on the projections from the current policies scenario from Keramidas and others (2024). In 2030, this would result in the share of fossil fuels in total electricity generation to be 16 percent in China, 37 percent in the United States, and 42 percent in Europe. Natural gas is the dominant fossil-fuel power source in the United States and Europe, whereas coal is prominent in China. The share of power generation from oil is very low (less than 1 percent) in all regions. The specific characteristics of the IT sector under the baseline are as follows. From the GTAP v11 database, for the IT sector the global average input cost shares of labor and capital were about 30 percent in 2017. Among the intermediate inputs about a quarter come from other compute services followed by about 5 percent from manufacturing inputs. 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 broadly similar in the U.S., China and Europe. Recent data indicates that AI platforms and service companies have almost doubled their electricity cost share in less than five years, from 0.8% in 2019 to 1.5% in 2023 (See Figure 2, panel 2). We assume that the rising trend in IT's electricity intensity would continue in the U.S. and by 2030 it increases to 4 percent in the U.S. from 1 percent in 2017. For the other countries these shares are kept identical to the 2017 values.

In both AI scenarios, we implement an AI-driven total factor productivity (TFP) shock in the IT sector that is calibrated such that the sectoral electricity demand from the IT sector matches the forecasts shown in Table 1. The TFP shock is implemented on the VA bundle in the nested-CES production function, and the TFP shock is only applied to production that uses new capital. For the United States, the TFP increase assumes a conservative rise in IT sector electricity intensity from 1.2 to 4 percent, well below the 13–14 percent intensity seen in data centers (Figure 3), which can be considered an upper limit.

The first AI scenario, *AI scenario under current energy policies,* introduces this AI shock under current energy policies that are identical to the baseline, assuming unchanged electricity generation mix. This leads to electricity generation across various technologies increasing in proportion to the overall rise in electricity supply. The second AI scenario, *AI scenario under alternative energy policies*, incorporates additional energy policies, boosting renewables' share via feed-in tariffs aligned with regional long-term strategies following NDC-LTS (Keramidas and others (2024)). Until 2030, renewable share in total generation under the NDC-LTS scenario is only slightly higher than what the *AI scenario under current policies* would achieve in most regions. Thus, the *AI scenario under alternative energy* policies does not impose significant climate or energy policy tightening until 2030.

Governments can use demand-side (for example, carbon taxes) and supply-side policies to shift electricity generation to less carbon-intensive sources. The AI expansion challenge hinges on electricity supply growth, so energy policies should focus on the stimulating the supply side. Among the supply-side policies, we implement feed-in tariffs for renewable energy sources though there are alternate instruments like feebates and regulations that can be used, however, with different macroeconomic, price and emission impacts (Chateau et al. (2024)). This policy incentive has been historically utilized as part of their policy frameworks in all countries where the AI shock is modelled. Moreover, the targeted renewable technologies are among the least cost options available in these regions and are even more competitive than fossil fuels (IRENA (2024)).

Several factors could potentially slow the growth of solar and wind capacity in countries 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 required for updating and expanding the grid, potentially limiting the expansion of renewable energy. For both AI scenarios, the two AI scenarios are also simulated by changing assumptions concerning medium-term constraints to test the sensitivity of the results to these rigidities. These constraints are modelled in two additional scenarios- (1) Capping the growth potential of renewables between 2025-

2030 to the average growth rate seen in the last five years (*Current/Alternative policies with smaller renewables scaleup*), and (2) Limiting new investment in transmission and distribution infrastructure compared to *baseline* (*Current/Alternative policies with no additional investments in T&D*).

Technically, the first constraint is modelled in IMF-ENV by capping the growth rates of solar PV and wind power generation to the recent years. In IMF-ENV, all power generation expansion needs to be supported by complementary expansion of transmission and distribution (T&D) sector. The second constraint is modelled by fixing the sectoral investment pathway of the T&D sector to that in the baseline pathway. Sectoral investments in T&D sectors are increasing in the *baseline* scenario, so the constraint reflects a case where the sector's expansion is not adequately keeping pace with the additional increase in new power generation capacity that is added in the economy under the AI shock. In all scenarios, power generation from hydropower and nuclear technologies is capped to baseline generation levels as expansion of these technologies largely depends on political decisions and geographical capacity rather than market mechanisms. Though this could result in results where power generation from these technologies falls below the levels in the *baseline* scenario, implying lower utilization rates, as a response to market dynamics where cheaper electricity sources are added to the power capacity.

In the *AI scenario under current energy policies*, the AI shock increases global GDP growth by 0.5 percentage points annually from 2025 to 2030, consistent with previous IMF estimates of 0.1 to 0.8 percentage points (IMF 2024). Of the total GDP impact, approximately 65 percent stems from expansion in IT sector production of AI services, and the remaining 35 percent reflects growth in other sectors through increased AI adoption and macroeconomic adjustments, including changes in energy prices.

References

Aguiar, A., Chepeliev, M., Corong, E. & van der Mensbrugghe, D. (2022). The GTAP Data Base: Version 11. *Journal of Global Economic Analysis, 7(2)*, 1-37.

Chateau, J., Jaumotte, F. and Schwerhoff, G., 2024. Climate policy options: a comparison of economic performance. Energy Policy, 192, p.114232.

Chateau, J., Rojas-Romagosa, H., Thube, S. D., & van der Mensbrugghe, D. (forthcoming). "*IMF-ENV: Integrating climate, energy, and trade policies in a general equilibrium framework*. IMF.

Keramidas, Kimon, Florian Fosse, Javier Aycart Lazo, Paul Dowling, Rafael Garaffa, Jose Ordonez, Stefan Petrovic, and others. 2025. Global Energy and Climate Outlook 2024. Luxembourg: Publications Office of the European Union. .

IEA (2024), Electricity 2024, IEA, Paris https://www.iea.org/reports/electricity-2024, Licence: CC BY 4.0

International Monetary Fund. 2024. World Economic Outlook— Steady but Slow: Resilience amid Divergence. Washington, DC. April

IRENA (2024), Renewable power generation costs in 2023, International Renewable Energy Agency, Abu Dhabi

Shehabi, A., Smith, S.J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M.A.B., Holecek, B., Koomey, J., Masanet, E., Sartor, D. 2024. 2024 United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-2001637