Powering AI: How Do Data Centers Affect Renewable Energy Investment?

Alexander Heiss, Zacharias Sautner, Thomas Schmid*

March 2026

ABSTRACT

This paper examines how the AI revolution affects investment in electricity generation. Exploiting the public release of ChatGPT as a plausibly exogenous shock, we show that power plant investments increase substantially more in regions with data centers than in regions without them. In contrast to the common view that data center growth is primarily supported by gas-fired generation, most of the newly announced capacity comes from renewable power plants and battery storage. A key mechanism behind these effects is data centers' use of power purchase agreements, which enable project-financed renewable energy investments. Furthermore, net-zero climate commitments by hyperscalers act as additional catalysts for data center-related renewable energy investments. These findings suggest that the growing electricity demand from data centers does not necessarily undermine the net-zero transition.

*Alexander Heiss is at Technical University of Munich and Boston Consulting Group, Thomas Schmid (corresponding author, schmid@hku.hk) is at The University of Hong Kong, and Zacharias Sautner is at University of Zurich and Swiss Finance Institute. Sautner is on the Sustainability Council of Lampe Asset Management. We thank Tobias Gutewort for excellent research assistance.

I. Introduction

Data centers transfer electricity into information. As a result, the rapid expansion of AI-related data centers represents a sharp shock to the demand for electricity.[1] Whether this demand surge is primarily met through investments in renewables or instead through fossil-fuel-based electricity is an important question: if the enormous growth in data center-related power demand is met primarily with fossil fuels, it can adversely affect the economy's decarbonization.[2] Moreover, how economies respond to a shock in electricity demand offers a laboratory for understanding how the future path of decarbonization may look. This is important to understand as decarbonization requires large-scale electricity generation from clean energy sources, making investments in new low-carbon electricity generation capacity a central economic challenge. Changes in energy infrastructure due to AI-related data center growth can therefore provide insights into whether and how economies can support scalable, low-carbon electrification in the decades ahead.

The answer to this important question is hotly debated. One common view is that the increased electricity demand from data centers is met, at least in the short run, by gas-fired generation rather than by new renewable supply (Volcovici and Kearney, 2024; Muir, 2025). The argument is that data centers require highly reliable 24/7 electricity, and the intermittency of wind and solar—especially in the absence of sufficient storage capacity—leads to more electricity generation from fossil fuels. In regions where grid infrastructure is constrained, onsite gas-fired power generation may also be the only feasible option for data center operators. Furthermore, gas-fired power plants are relatively inexpensive and quick to build. Consequently, the sharply increasing energy demand of data centers may reinforce reliance on conventional power sources instead of new investments in renewables.

An alternative view holds that data centers spur renewable energy investment because their large and predictable electricity demand creates strong incentives to build new clean energy capacity. The argument is that hyperscalers such as Google, Microsoft, or Amazon have made public commitments to powering their data centers with renewables and sign long-term power purchase agreements (PPAs) with renewables developers to satisfy their energy demands. PPA contracts are more relevant for renewable projects than for gas plants because the economics of revenue risk and operating flexibility differ across these

[1] According to recent estimates, up to 50% of the growth in electricity demand in the United States until 2030 originates from data centers (International Energy Agency, 2025).

[2] Beyond this potential effect, the deployment of AI tools that run on data centers may help reduce emissions elsewhere in the economy; this effect needs to be accounted for when evaluating the overall environmental effects of AI.

technologies (Darmouni et al., 2026).[3] These contracts reduce financial risk and facilitate financing for new renewable projects. In addition, the presence of large data centers can encourage upgrades to grid infrastructure and energy storage, enabling greater integration of renewable sources into electricity systems. As a result, the rapid growth of data centers can accelerate investments in renewables.

Convincingly answering which of these two views holds true in the data is difficult. On the identification side, a challenge is that data centers are not randomly located. Operators often choose locations with abundant electricity supply, low prices, strong grid infrastructure, or existing renewable resources. These factors also influence renewable investment decisions. As a result, correlations between data centers and renewable investments may reflect underlying regional characteristics rather than causal effects. Relatedly, developers may invest in power capacity before data centers are actually built, anticipating future demand, making it hard to identify whether renewable energy investments are a response to data center demand. Renewable energy investments might also be influenced by policy changes, and regions attracting data centers may simultaneously implement policies to encourage clean energy. On the data side, a challenge is to obtain detailed information on data centers—especially planned ones—their electricity demand, and locations. Additionally, one needs granular information on actual and planned investments in power generation capacity, the respective locations, how they are financed, and how the generated electricity is sold.

We make substantial progress in addressing each of these challenges, providing the first large-scale causal evidence on how the AI revolution affects investments in different power generation technologies in the United States.[4] Our identification strategy exploits the unexpected surge in future electricity demand from data centers following the public release of ChatGPT in November 2022 as an exogenous shock in a difference-in-differences (DiD) design. We divide the continental U.S. into uniform 100×100 km grid cells, and determine treatment status using *pre-shock* data center capacity: cells in the top 10% of capacity are classified as treated, accounting for 92.9% of the pre-ChatGPT data center capacity, while others serve as controls. Because data centers and power plants cluster locally due to grid constraints, transmission costs, and the design of U.S. electricity markets, changes in electricity demand primarily affect nearby power investment decisions. Exploiting these institutional frictions, we compare changes in investment in different

[3] In the data we use, 47% of the investments in renewables capacity use PPAs, while the corresponding number for gas-fired capacity is only 6%.

[4] Around 40% of the global electricity demand related to data centers comes from the U.S. (International Energy Agency, 2025).

power plant technologies before and after the ChatGPT shock between data center and non-data center regions. Because treatment is defined using pre-shock data center locations, and because the release of ChatGPT generated an exogenous increase in expected electricity demand from data centers, the DiD specification identifies the causal effect of data center-driven electricity demand on power plant investment. We additionally control for region and time fixed effects, local electricity system characteristics, and policy shocks at the state level.

We address the data challenge by collecting rich project-level data from the U.S. Energy Information Administration (EIA) on existing and planned power generating units and storage facilities. We complement these data with hand-collected information on behind-the-meter (BTM) power plants, which are co-located with data centers and not covered by the EIA data as they are not connected to the grid. We compile information on existing and planned data centers from the S&P 451 Research's Datacenter KnowledgeBase, which provides facility-level information on the start of operation, estimated power capacity, developer or operator identity, and exact location.

We show that planned power plant capacity increases substantially more in regions with existing, pre-ChatGPT data centers than in regions without them after the public release of ChatGPT. In the post-treatment period, treated regions are 8.2 percentage points more likely to experience a new power plant announcement, corresponding to an average capacity increase of 30 megawatts (MW) per quarter. On the intensive margin, the estimated treatment effect remains positive but is not statistically significant. This pattern suggests that the aggregate treatment effect operates primarily through the extensive margin: treated regions become more likely to see new power plant announcements, rather than experiencing an increase in the size of individual projects.

Crucially, most of this new capacity comes from renewable power plants and battery energy storage systems (BESS) rather than gas-fired plants.[5] For renewables and BESS, we find additional capacity investments both at the extensive and intensive margin. When we decompose investments for renewables into wind, solar, and BESS, and for has into grid-connected and BTM plants, the treatment effect in data center regions is driven by increased investments in solar and BESS, with no effect for wind. Not surprisingly, we observe a small increase in BTM gas plants in data center regions, but this increase is

[5]BESS are installations that store electricity in batteries—typically lithium-ion—to save excess power and release it later. The source of power is typically renewables. Renewable power plants and BESS are technologically complementary and frequently co-located. The complementarity arises from the ability of BESS to mitigate the intermittency of wind and solar generation by storing excess electricity and discharging it when renewables output is low.

relatively small and largely offset by a decline in grid-connected gas plants, causing the net effect to be zero. These findings indicate that the AI boom increases power plant investment in treated regions, and this increase is concentrated in renewables and BESS rather than in gas-fired generation. Our results challenge the narrative that the growth in electricity demand from data centers is largely filled with fossil fuels.

To strengthen the credibility of our findings, we conduct additional analyses and robustness tests. First, when we estimate dynamic treatment effects around the 2022 treatment year, we find no evidence of pre-trends, while significant positive effects emerge especially in 2024–2025—driven by renewable plants and BESS but not gas plants. Second, replacing the binary treatment with treatment deciles shows that these investment effects are strongest in regions with the highest pre-existing data center capacity, indicating that the effect increases with exposure to data centers. Third, when we include state-by-time fixed effects to account for policy shocks, ensuring identification comes from within-state variation, we continue to obtain our results. Fourth, robustness checks altering treatment definitions, timing windows, and grid-cell sizes yield consistent results.

To understand the economics underlying these results, we document that data centers' use of PPAs is a key mechanism behind the investment effects. We show that PPAs and project finance serve as channels linking the AI-driven demand shock to higher investments in renewables. Using plant-level data, we find that power plants near data centers became more likely to use PPAs and project finance after the release of ChatGPT, suggesting that data center demand expands financing opportunities for renewable projects. Additional tests confirm that PPAs and project finance are significantly more common for renewable plants and BESS than for other technologies, supporting the idea that the AI boom stimulates renewable investment through improved financing conditions.

These results are plausible. The reason is that volatile electricity prices create strong incentives to secure long-term electricity supply through PPAs.[6] PPAs also allow large buyers to contract electricity at predetermined prices, reducing price risk and procurement uncertainty while also facilitating the financing of power plants, since stable long-term revenues support project finance structures that rely on predictable cash flows. This mechanism is particularly important for renewable power plants, whose revenues are relatively more exposed to electricity price volatility and therefore benefit more from PPAs that stabilize income, increase feasible leverage, and lower financing costs; in contrast,

[6]Data center operators are attractive PPA counterparties because their high credit quality mitigates default risk, increasing the safety of the contracted revenue stream for the power plant operator. The four largest data center operators in our sample—Amazon, Meta, Microsoft, and Alphabet—all have AA- or better S&P credit ratings as of March 2026.

gas plants rely less on such contracts (Darmouni et al., 2026).

Finally, we document that prior net-zero pledges by data center operators act as a commitment mechanism that facilitates this effect. Using a data center–level analysis, we find that the increase in renewable investments is concentrated around data centers whose operators have climate pledges, while this is not the case for gas plants. This suggests that climate commitments are credible and shape both the amount of investment in electricity capacity around data centers and its technology composition, steering it toward cleaner sources. These results are also plausible. For large technology firms, especially hyperscalers such as Amazon, Microsoft, or Google, failing to meet publicly announced climate targets would expose them to reputational costs, investor scrutiny, and potential regulatory pressure. Moreover, the long-term planning horizons of data center investments and their strong financial performance allow technology firms to focus their electricity procurement strategies on meeting climate targets. As a result, renewables developers may interpret such commitments as credible signals of persistent future demand for low-carbon electricity, which can improve expectations about the profitability of renewable generation projects and stimulate investment in new renewable capacity.

Our findings contribute to different strands of the literature. Most directly, we relate to a nascent literature examining the implications of data centers on power generation and climate change. This work comes from industry studies (e.g., International Energy Agency, 2025) and lacks credible identification and an analysis of economic mechanisms (e.g., Guidi et al., 2024). Some recent work estimates the effects of data centers on electricity prices and market efficiency (e.g., Feher et al., 2025; Mamkhezri et al., 2025; Kay et al., 2026). We further relate to Heiss and Schmid (2026), who study how project-financed power plants use PPAs to hedge revenue risk, and Darmouni et al. (2026), who show that PPAs mitigate price risk for renewable electricity, enabling renewable investment through a capital-structure channel. Indirectly, we contribute to work examining the importance of climate commitments for firms' investment in decarbonization (Bolton and Kacperczyk, 2026; Bauer et al., 2024). Our results show that green pledges act as credible commitments to invest in renewables in a setting that would otherwise be very carbon intensive.

II. Institutional Setting and Predictions

This section describes the link between data centers and investments in renewables, explaining the location choice of data centers, their electricity demand, and the spatial

clustering of data centers and power plants. It then explains the importance of PPAs and climate pledges by data center operators for investments in renewables. The discussed institutional features allow us to define the identification strategy, derive predictions, and isolate economic channels underlying our results.

A. Data Centers and Investments in Renewables

A.1. Data Center Location Decisions

Hyperscale data centers are among the most electricity-intensive commercial facilities and exhibit high load factors (Masanet et al., 2020). Economically, they resemble baseload industrial consumption: once built, utilization is relatively steady, power outages are extremely costly, and short-run demand elasticity is limited (Brancucci et al., 2025). As a consequence, developers place high value on the reliability and redundancy of electricity supply, timely interconnection to the electricity grid, low and predictable electricity costs, and supporting infrastructure such as fiber connectivity, land availability, and permitting capacity. Importantly, these factors are complementary: locations that already host large digital infrastructure tend to attract additional investment because incremental builds can leverage, among others, existing electricity and fiber network infrastructure. Data centers typically need 1 to 1.5 years to be built, so electricity needs to be available quickly.

Figure 1 illustrates the spatial distribution of existing and planned data centers as of 2022Q3 in Subfigure (a) (before the release of ChatGPT) and as of 2025Q3 in Subfigure (b) (after ChatGPT), calculated based on data that we explain in the next section. The figures show a strong clustering of data centers, mainly in Northern Virginia and Texas, and a high correlation of data center locations over time. Northern Virginia has emerged as the leading U.S. hub, in part due to dense fiber infrastructure and proximity to major network exchange points, the availability of reliable and comparatively inexpensive electricity, and tax incentives (Virginia JLARC, 2025). Texas has similarly attracted large-scale data center development, in part through a deregulated power market with cheap energy, rapid renewable power expansion options, and state-level sales and use tax exemptions for data centers (Texas Comptroller of Public Accounts, 2024).

Conversely, other states like Florida or California, with large populations and high economic activity, host relatively few hyperscale data centers. A key reason is that siting is constrained by the ability to secure timely and reliable electricity interconnections and by high electricity costs (Mytton et al., 2023; Brancucci et al., 2025). These factors make locations less attractive for hyperscale builds, particularly when developers can substitute

toward other regions. Furthermore, exposure to natural hazards can affect both insurance costs and expected downtime risk, leading to the avoidance of high-risk regions such as California. The clustering of data centers is also consistent with standard agglomeration mechanisms, including shared infrastructure, specialized labor and suppliers, and coordination or learning spillovers (Ellison and Glaeser, 1997).

A.2. Data Center Power Demand and the ChatGPT Shock

As hyperscale data centers constitute highly electricity-intensive infrastructure, with many facilities designed to operate at power capacities in excess of 100 MW—data centers are therefore also typically classified based on their respective power needs, not the computing technology used. For example, Alphabet's Red Oak 1 hyperscale data center in Texas, which came online in 2025Q1, has a power capacity of 270 MW, roughly equivalent to the electricity demand of 180,000 U.S. residents. Data centers can source their electricity from grid-connected power plants or non-grid-connected BTM generation plants. BTM plants virtually always operate with gas turbines.

Figure 2 shows that the electricity demand of U.S. data centers has risen from 108 TWh in 2020 to 183 TWh in 2024, and is expected to grow to 426 TWh by 2030 in the base case scenario (International Energy Agency, 2025). The data center share of electricity demand, relative to total U.S. power consumption, has increased by over 50% from 2020 to 2024 (from 2.9% to 4.6%), and is expected to more than double until 2030. Thus, the rapid expansion of data centers creates a large and geographically concentrated electricity demand. As a result, after years of relatively stable electricity demand, data centers have emerged as the key driver of growth in electricity demand.[7]

Expectations about the growth in electricity demand by data centers were shifted substantially by the release of ChatGPT by OpenAI on 30 November 2022. This event revealed that large language models can scale to the mass market and led to a rapid reassessment of the electricity required by data centers that train and host these models. Consistent with a significant shift in expectations, grid planners and utilities sharply revised their load forecasts upward after ChatGPT's release. For example, utility forecasts submitted to the Federal Energy Regulatory Commission (FERC) show that the projected U.S.-wide summer peak load growth over the next five years rose from 24 Gigawatts (GW) in 2022 before the ChatGPT shock to 38 GW in 2023, 64 GW in 2024, and 166 GW in

[7]Net electricity generation in the U.S. increased only modestly in the two decades before the release of ChatGPT, from 3,858 TWh in 2002 to 4,065 TWh in 2012 and 4,183 TWh in 2022.

2025, with data centers accounting for the majority of the growth (FERC, 2026).[8]

Equity market reactions confirm that the release of ChatGPT constitutes an unexpected shock to data center expansions. NVIDIA, the key supplier of GPUs used in hyperscale data centers, had a relatively flat stock price development before the release of ChatGPT (from $13.33 on November 1, 2020 to $13.50 on October 31, 2022). In the two years following ChatGPT's release, the stock price of NVIDIA rose from $17.14 to $138.20 as investors revised their expectations about the future demand for data center growth and their GPUs.

Thus, energy demand forecast revisions and stock market reactions indicate that the ChatGPT release constitutes a shock that sharply lifted expectations about the level and growth rate of data center electricity demand. Reflecting the unexpected, material nature of the event, the ChatGPT shock is also used in other settings (e.g., Hui et al., 2023).

A.3. Co-Location of Data Centers and Power Plants

Large U.S. data centers typically trigger the construction of new power plants and storage projects *in the same geographic area* for three related reasons. First, transmission networks have hard capacity limits. Lines, transformers, and substations are often congested, so bringing a large, concentrated new load into a network is typically infeasible without major upgrades. These upgrades take years to permit and build (Hogan, 1992; U.S. Department of Energy, Grid Deployment Office, 2023). Building new generation capacity close to data centers makes it easier to build direct connections and reduces a project's reliance on constrained grid connections.

Second, electricity transmission is constrained by physics due to transmission losses that increase with distance. When power is transmitted farther, more energy is lost as heat on the wires, implying that more electricity generation is necessary to deliver the same amount of power to a data center (Eldridge et al., 2017). Additionally, electricity is usually stepped up to higher voltage levels for transmission over medium or long distances, which causes further losses. Thus, the distance between a data center and the energy source increases electricity costs, making nearby generation more efficient and cheaper.[9]

Third, the design of U.S. electricity markets incentivizes local siting of electricity

[8]Regional planning documents similarly highlight large load adjustments tied to data centers. An example is PJM, the Regional Transmission Organization (RTO) that manages parts of the U.S. grid, including Virginia. In 2023, right after the ChatGPT release, PJM massively increased their load forecast due to the additional demand for data centers (PJM Interconnection, 2023).

[9]Several power plant developers state that they chose the plant location because of nearby power demand from AI data centers, such as Chevron or El Paso Electric.

generation. Organized wholesale markets use nodal locational marginal pricing (LMP), which sets a price that reflects the marginal cost of serving load at each grid node (ISO New England, 2026). This price-setting is very granular, with approximately 60,000 pricing nodes. LMPs can differ substantially even within small areas because the LMP does not only include the electricity generation cost, but also local congestion and transmission losses (Hogan, 1992; Borenstein et al., 2002).[10] When congestion is high near a data center hub, local prices rise relative to surrounding nodes, so generation and storage that interconnect nearby earn higher revenues and provide more reliable supply.[11]

To summarize, we predict that power plant capacity increases substantially more in regions with data centers than in regions without them. In the next subsection, we predict that much of this increase is met by investments in renewable energy.

B. Catalysts for Investments in Renewables

B.1. Project Finance and Power Purchase Agreements

For data center operators, electricity is both a major operating cost and a critical input for their operations. However, electricity prices in deregulated markets are highly volatile (Lin et al., 2021), and the massive power demand from data centers can create supply constraints. These factors create strong incentives for data center operators to secure their long-term electricity supply at predictable prices. This goal can be achieved through PPAs, which enable large buyers to contract directly with specific power plants. PPAs are contracts between consumers and generators to deliver electricity at future dates for pre-determined prices. Thus, PPAs hedge price risk for data center operators.[12]

PPAs matter for the financing of power plants as many utility-scale plants are project-financed as special-purpose vehicles with limited or non-recourse debt (in our sample, 35% of power plants, representing 40% of total capacity, are project-financed). In the project finance model, lenders are paid from project cash flows, and debt capacity depends on the stability of those cash flows (Esty and Megginson, 2003). Long-term revenue hedging

[10] For instance, in Iowa the LMP on September 7, 2011 ranged from negative prices to more than $100 per MWh (see EIA 2011).

[11] In contrast, most European power markets use zonal pricing, where a single price applies to a country or region. This reduces the incentive to build power plants near power consumers (ACER, 2026)).

[12] There are two types of PPAs: Utility PPAs are long-term contracts in which an electricity utility agrees to buy power from a generator at a predetermined price and conditions. Corporate PPAs, now legalized in most larger U.S. states as a result of the deregulation of power markets, allow independent power producers (IPPs) to sell electricity directly to corporates without the involvement of utilities (for details, see Heiss and Schmid, 2026).

is central for convincing debt providers to lend to the project. PPAs achieve this purpose because they allow generators to hedge prices and quantities over long horizons (often more than ten years).[13] Consistent with this argument, Heiss and Schmid (2026) show that most project-financed power plants use PPAs to hedge their future revenues.

PPAs and project finance are not equally important for all power plant types. Rather, they are more relevant for renewable plants than for gas plants because the economics of revenue risk and operating flexibility differ across these technologies. This is shown by Darmouni et al. (2026), who develop and calibrate a framework in which renewables' near-zero marginal costs make cash flows especially sensitive to electricity price risk. PPAs reduce this risk, increase feasible leverage in project finance, lower discount rates, and raise renewable investments through a capital structure channel. In contrast, gas plants face volatile input costs and create value from dispatch flexibility, which makes locking in price and quantity via PPAs less attractive for them. In our sample, 47% of renewable generation capacity use a PPA, while the corresponding number for gas-fired capacity is only 6%. The importance of PPAs for renewable plants is summarized in a quote from Patrick Pfeiffer, Head of Statkraft U.S., an electricity trader and market operator:

> *Renewable energy PPAs provide a pathway to a net zero future. They are essential for generators, end users, and society as a whole. For most generators, securing bank financing hinges on having a credit-worthy buyer for the energy they will produce.* (Statkraft, March 26, 2024)

The volume that traditional PPA counterparties, such as utilities, are willing to contract with power generators does not keep pace with the growth of renewables. As a result, corporations are increasingly stepping in to sign PPAs with developers. According to Bloomberg, the global market for such Corporate PPAs expanded from less than 5 GW in 2015 to nearly 60 GW in 2025 (BloombergNEF, 2026). The four largest buyers of electricity via Corporate PPAs are all data center operators: Amazon, Meta, Alphabet, and Microsoft. Consequently, these contracts provide renewables developers with a critical—and often the only—opportunity to hedge their future revenue risk.

As reflected in the above quote, it is vital for renewables developers to find a counterparty with high creditworthiness given the long duration of PPAs.[14] Data center operators

[13] This logic is similar to corporate finance theories in which risk management can be value-enhancing when external finance is costly and cash flow volatility has real effects on investment (see Froot et al., 1993; Campello et al., 2011). Empirical evidence shows that access to risk management tools can affect real outcomes and firm value in settings where cash flow risk interacts with financing frictions (Pérez-González and Yun, 2013; Cornaggia, 2013; Gilje and Taillard, 2017).

[14] While default risk was a minimal concern for traditional utility PPAs, it is more pronounced for

are attractive counterparties because their high credit quality mitigates default risk, increasing the safety of the revenue stream for the power plant operator. The four largest data center operators in our sample—Amazon, Meta, Microsoft, and Alphabet—all have AA- or better S&P credit ratings as of March 2026. As a consequence, the growth of data center load not only improves the quantity of long-term contracting demand but also the quality of the available counterparties for power plant operators.[15] In this context, Forbes Magazine states that "[t]he data center industry is a major buyer of PPA agreements for RE [...] helping to drive decarbonization by underwriting a significant proportion of grid-scale, carbon-free energy for industry."

To summarize, we predict that the PPA demand by data center operators creates a financing mechanism through which the AI boom affects power plant investments, especially for renewables.

B.2. Climate Commitments

Net-zero commitments by data center operators can function as an additional catalyst for renewable investments, because these commitments are often perceived as credible and likely to be fulfilled rather than merely symbolic. The reason is that for large technology firms, especially the hyperscalers, failing to meet publicly announced climate targets would expose them to reputational costs, investor scrutiny, and potential regulatory pressure, particularly as environmental performance has become increasingly material for capital markets.[16] Moreover, the long-term planning horizons of data center investments and their strong financial performance allow technology firms to focus their electricity procurement strategies on meeting climate targets. Renewables developers may interpret such commitments as credible signals of persistent future demand for low-carbon electricity, which can improve expectations about the profitability of renewable generation projects and stimulate investment. Evidence consistent with the argument that climate commitments are credible is provided by Bauer et al. (2024) and Bolton and Kacperczyk (2026). These papers show that firms that make green pledges subsequently reduce their carbon emissions and the announcements of these pledges persistently raise stock prices.

These arguments lead to the prediction that climate commitments by data center

Corporate PPAs. A report by Baker McKenzie in 2015 states that "[t]he core issue for banks is that most corporate off-takers are likely to have a much worse credit rating than a load-serving utility."

[15]Empirical evidence highlights the importance of contracting frictions and enforcement risk for creditor exposure (e.g., Esty and Megginson, 2003), and work on green energy procurement shows that counterparty risk can raise contract prices and depress investment (e.g., Ryan, 2021).

[16]Google, for example, aims to achieve net-zero emissions by 2030 covering Scope 1, 2, and 3 emissions.

operators catalyze investments in renewable energy.

III. Data

This section explains how we source the data on power plants and data centers. Additionally, we describe the construction of key variables. Table 1 reports summary statistics.

A. *Data on Data Centers*

We obtain information on data centers from the S&P 451 Research's Datacenter KnowledgeBase (S&P 451 in short), which provides the most established and comprehensive overview of data centers.[17] Information on data centers is compiled by S&P 451 based on briefings and surveys with data center owners and end users (S&P Global, 2026). The database provides facility-level information on the start of operation, power capacity (in MW), developer or operator identity, and location (in the form of latitude and longitude). Included in the database are existing as well as planned data centers. For planned data centers, the database contains information on the anticipated date of when they will become operational. The most recent version of the data at the time of writing is 2025Q3. In addition to this 2025Q3 version, we use a historical snapshot of the same database from 2022Q3, predating the release of ChatGPT. This snapshot is critical to define data center regions: to avoid concerns about reverse causality, using the pre-ChatGPT data excludes the possibility that the siting of new data centers is based on pre-existing plans to construct new power plants in a region (see Section IV for details).

Figure 3 plots data center power capacity by operating year from 2000 to 2032 using the 2025Q3 data version. The figure shows an increasing trend in power capacity since approximately 2010, which accelerated substantially after 2022.[18]

B. *Data on Power Plant Investments*

To measure power plant investments, we obtain project-level data from the U.S. Energy Information Administration's (EIA) Form 860M filings. These filings report existing and planned power generating units and storage facilities. The data include information on

[17] S&P 451 data is commonly used in industry reports (e.g., International Energy Agency, 2025) and in recent academic research (e.g., Feher et al., 2025; Brannon and Wolf, 2024).

[18] The post-2028 decline reflects that most data centers that will become operational by then are not yet announced.

the plants' technologies, locations (latitude and longitude), and actual or planned start of operation. The filings do not only cover existing power plants, but also those that are planned or under construction. These data are available on a monthly level from 2016 to 2025, which also defines our sample period. BTM power plants are not covered in Form 860M filings and to collect information on them, we manually compiled BTM plant data related to the data centers in our sample.[19] We measure power plant investments by the capacity of newly announced plants. A plant is "announced" in the month in which it first appears with a status of "planned" in Form 860M or, for BTM generation, in which its existence is first reported. These monthly announced investments are then aggregated to the year-quarter or year level, depending on the analysis.

Figure 4 plots announced power plant capacity from 2016 to 2025 (in MW per year). We also report the underlying technologies, that is, wind, solar, gas, and BESS, and for gas whether it is grid or BTM production. There is a modest increase in announced capacity between 2016 and 2023, a substantial jump from 2023 to 2024, and a slight decrease—at a high level—from 2024 to 2025. The jump in 2024 is driven by an increase in solar and BESS capacity and, at a lower level, by gas capacity. BTM gas generation played a minor role until 2024 but became more common in 2025.

Figure 5 shows the spatial distribution of announced power plants by technology. New plants were announced in most areas of the U.S. during our sample period. As would be expected given the need to produce power close to where demand locates, there is a clustering of new plants in the populous parts of the U.S. (e.g., in California, Texas, or along the East Coast). Below we illustrate that the spatial and technological distribution of newly announced capacity strongly relates to the electricity demand by data centers.

C. Data on PPAs, Financing Structures, and Climate Commitments

We match power plants to the S&P Capital IQ Pro database to obtain information on owners and contractual arrangements, including indicators for the presence of PPAs. To obtain information on the financing structures of power plants, we link plants to LSEG Refinitiv's project finance database, allowing us to distinguish project finance from corporate financing.[20] We only keep plants with a capacity of more than 10 MW as the coverage of project finance deals for smaller power plants is limited.

[19] We use the "Deep Research modes" of Google Gemini and ChatGPT to check for government approvals, press releases, news articles, or similar sources that describe how a data center is powered.

[20] The matching is done based on name, owner name, power plant capacity, fuel type, and location. For further details on the matching process see Heiss and Schmid (2026).

We hand-collect data on climate commitments of data center operators. Specifically, we classify whether a data center operator has publicly stated an ambition consistent with a carbon net-zero target or equivalent climate pledge.[21] We then create an indicator for operators with such a pledge and assign it to all of their data centers.

IV. Methodology

Our methodology exploits that the public release of ChatGPT led to a massive and unexpected increase in expected electricity demand from data centers. We use this shock as the time treatment in a DiD design. Cross-sectionally, we exploit the spatial clustering and co-location of data centers and power plants, which occurs because of grid capacity limits, transmission losses, and price-setting at the granular nodal level (see Section II). This institutional feature enables us to define local regions and analyze data centers and power plant investments within those regions before and after the ChatGPT shock. To ensure that all regions are granular and of equal size, we define them using a uniform spatial grid rather than administrative borders: we partition the continental U.S. into approximately 100×100 km grid cells.[22] We exclude grid cells with less than one MW of existing capacity and use various larger or smaller grid sizes in robustness tests to make sure that the choice of a 100 km grid does not drive our findings.

Crucially, to differentiate treated and control regions, we rely on the presence of data centers *before* the release of ChatGPT. We define those with high data center capacity before ChatGPT as treated regions and those with low or no capacity as control regions. For this purpose, we use the 2022Q3 snapshot of the S&P 451 database, in order to identify only data centers that were either operational or planned *before* ChatGPT was released. This approach mitigates concerns about reverse causality: it excludes the possibility that the siting of new data centers is based on pre-existing plans to construct new power plants. The assumption behind our approach is that historical data center regions are meaningful predictors for the siting of data centers after the treatment, which is supported by the spatial clustering of data centers (see Section II.A.1). We empirically validate this assumption in Appendix B: the table shows a strong correlation between regions where

[21] We obtain these data by using the "Deep Research" modes of OpenAI's ChatGPT and Google Gemini to search for a publicly stated net-zero commitment year for each individual operator. In cases where the two models obtain different results, we manually check the sources like websites and press announcements.

[22] The grid is implemented in latitude–longitude degrees (0.90° latitude by 1.16° longitude, which is approximately 100 km at 39°N).

data centers were located or announced before the ChatGPT shock and those that were announced afterwards.[23]

To measure data center capacity in a grid cell before the release of ChatGPT, we use all existing and planned data centers in the 2022Q3 snapshot of the S&P 451 database. We then rank grid cells by the capacity of data centers, ignoring very small data centers with less than 10 MW, and classify a region as treated if it ranks in the top 10% of grid cells over the U.S. mainland. Figure 6 shows the capacity of data centers as of 2022Q3 for each grid cell. A total of 73 cells are classified as treated. These cells are colored in green and account for 92.9% of the total data center capacity.

Using these classifications, we estimate variants of the following DiD regression:

$$Capa_{r,t} = \alpha + \beta Post_t \times Treat_r + \gamma \times \boldsymbol{C}_{r,t} + \rho_r + \tau_t + \epsilon_{r,t}, \tag{1}$$

where $Capa_{r,t}$ measures power plant investments in region r and year-quarter t, $Post_t$ equals one for year-quarters starting in 2022Q4 and zero before, and $Treat_r$ equals one for treated regions and zero for control regions. $\boldsymbol{C}_{r,t}$ is a vector of region-time specific control variables, ρ_r indicates region-fixed effects, and τ_t year-quarter fixed effects. The individual effects for $Post_t$ and $Treat_r$ are absorbed by these fixed effects. In an additional test, we include state times year-quarter fixed effects to ensure that the estimate for β is not biased by state-level policy shocks.

We construct different variables to measure newly announced power plant capacity in a region and quarter: (i) $Capa_{r,t}^{0/1}$ equals one if any new power plant is announced in year-quarter t and region r, and (ii) $Capa_{r,t}^{GW}$ measures the capacity of all newly announced power plants at time t in region r. For the latter variable, we also use log and log-plus-one versions.[24] In addition to these measures that reflect all newly announced power plants, we construct versions that only consider renewable power plants (together with BESS) or only gas-fired power plants. We consider renewable power plants and BESS jointly in this analysis because they are technologically complementary and frequently co-located. The complementarity arises from the ability of BESS to mitigate the intermittency of wind and solar generation by storing excess electricity and discharging it when renewables output is low. In further tests, we break down these variables into wind, solar, BESS,

[23] If power capacity is missing for planned power plants in the 2022Q3 snapshot, we complement missing capacity data with the power capacity from the 2025Q3 version. To ensure that this does not introduce a forward-looking bias, we also report results without this adjustment in the robustness section.

[24] We report results using the log-plus-one transformation, alongside the other dependent variables, despite its known methodological limitations (Chen and Roth, 2024).

grid-connected gas plants and BTM gas plants.

To account for region–time specific electricity system conditions, we include two control variables. First, $Log(Plant\,capa_{r,t})$ measures the logarithm of the total installed capacity of all existing power plants in region r at time t. This variable captures differences in the scale of regional electricity systems and serves as a proxy for underlying electricity demand and market size. Regions with larger installed capacity typically operate larger electricity markets and may exhibit systematically different investment patterns. Second, $Share\,RE/BESS_{r,t}$ measures the share of renewable energy and BESS capacity in total installed capacity in region r at time t. This variable captures the technological composition of the regional power generation mix and the progress of the energy transition. Since the penetration of renewables and storage can affect both system flexibility requirements and incentives for new generation investment, controlling for this share accounts for differences in regional electricity market structure and policy environments over time.

In addition to the main DiD analysis on the region level, we conduct DiD regressions at the data center and power plant level. At the data center level, we analyze new power plants that became operational within a radius of 10 km around the data center in the same year or one year before/after a data center's opening. At the power plant level, we focus on whether a power plant applies a PPA to hedge price risk and whether it is project or corporate financed. More details are provided in Section V.

V. Empirical Results

This section first shows graphical trends in the changes of announced power plant capacity in treated and control regions over time. After that, we show our main regression results. We then test the crucial DiD assumptions and document the robustness of the results. Finally, we examine the economic mechanisms driving the results.

A. Main Results: Graphical Illustration

Figure 7 graphically illustrates the main result of our paper. We plot the evolution of announced power plant capacity for treated (data center) and control (non-data center) regions over time. Subfigure (a) includes all power plants, Subfigure (b) includes only renewable plants and BESS, and Subfigure (c) only gas-fired power plants including both grid-connected and BTM facilities. The horizontal axis reports year-quarters and the vertical axis the cumulative capacity of the announced power plants since 2016Q1. For

each year-quarter, we aggregate newly announced capacity within a region and calculate averages separately for treated and control regions. We then normalize cumulative capacity for each region to 100% in 2022Q3 (just before ChatGPT was released). We winsorize the resulting variable at the 1^{st} and 99^{th} percentiles to limit the influence of outliers.

Subfigure (a) shows that cumulative announced capacity increases substantially faster in treated data center regions after the release of ChatGPT in 2022Q4. Control regions also exhibit some acceleration, but the increase is considerably less pronounced. Prior to 2022Q4, there is no visual difference in capacity growth across the two types of regions. The disaggregation by technology in Subfigure (b) and (c) indicates that the aggregate trend after the ChatGPT shock in Subfigure (a) is driven primarily by renewable plants and BESS and data center regions are largely responsible for this growth. By contrast, for gas-fired power plants, we find little evidence of differential investment trends between treated and control regions after the release of ChatGPT, albeit a small difference arises. Taken together, these patterns provide initial evidence that the AI boom spurred power plant investment, particularly in renewables and BESS.

Figure 8 adopts a slightly different aggregation procedure to illustrate the robustness of our finding. Instead of first constructing cumulative announced capacity at the regional level and then averaging across regions, we sum cumulative announced capacity directly across treated and control regions. In other words, while the previous figure assigned equal weight to each region, this specification weights regions by announced capacity. The results are similar: we continue to observe an increase in total announced capacity that is driven by renewable plants and BESS. However, the effects are less pronounced than in the previous figure, suggesting that some of the earlier patterns reflect very large power plant investments in a subset of regions.

B. Main Results: Regression Evidence

We next turn to more formal regression analysis and estimate Equation 1 at the region-quarter level. The results are reported in Table 2. Panel A uses all newly announced power plants in a given year-quarter and region, Panel B considers renewable plants and BESS, and Panel C focuses on gas-fired power plants. In each panel and for each technology type, the dependent variable in Column 1 is $Capa_{r,t}^{0/1}$ (equals one if any new plant is announced in year-quarter t and region r), in Column 2 it is $Capa_{r,t}^{GW}$ (the capacity of newly announced plants), and in Columns 3 and 4 it is the logarithms of $Capa_{r,t}$ and one plus $Capa_{r,t}$, respectively (we use capacity in MW for the log-versions of the measure).

Confirming the graphical evidence, power plant investment increases significantly more in treated regions than in control regions following the ChatGPT shock. In the baseline specifications in Panel A, the estimated treatment effect (β in Equation 1) is positive and statistically significant when the dependent variable captures the incidence of newly announced investment in Column 1; this indicates a clear post-treatment divergence between treated and control regions. Treated regions are 8.2 percentage points more likely to receive any new plant investment after the treatment, relative to the change in control regions. For $Capa_{r,t}$ in Column 2, the estimates imply that treated regions announced, on average, 0.03 GW (or 30 MW) more new power plant capacity in quarters after the treatment than would be expected based on the change in control regions.

By contrast, when we use $Log(Capa_{r,t})$ as the dependent variable, thereby focusing on the intensive margin (as region-quarters without new capacity are dropped by construction), the treatment coefficient remains positive but is not statistically significant. This suggests that the aggregate treatment effect operates through the extensive margin: treated regions become more likely to experience new power plant announcements, rather than these regions seeing an increase in the size of individual projects after the public release of ChatGPT. The estimates in Column 4, which identify effects again across all region-quarters, confirm those in Column 2, with treated regions increasing announced capacity by more than 50% relative to the counterfactual change in control regions.

Notably, Panel B confirms the graphical evidence that the overall increase in power plant investment is driven primarily by renewable plants and BESS. For these plants, the estimated effects are consistently positive and economically meaningful, at both the extensive and intensive margin, indicating that the post-ChatGPT increase in announced investment is concentrated in renewables and storage technologies. A comparison with the estimated coefficient in Panel A shows that most of the economic effects of the aggregate treatment originate from investments in renewables and BESS. In Column 1, treated regions are 7.5 percentage points more likely to receive any new renewables and BESS investment after the treatment, relative to the change in control regions. The treatment effect is now not confined to the extensive margin. In Column 3, we find also an increase in the size of renewables investments in data center regions. By contrast, the treatment effect for gas-fired plants in Panel C is much smaller economically, changes sign across specifications, and is statistically insignificant.

Overall, these results reinforce that the recent AI boom increases announced power plant investment in treated regions, and this increase originates mostly from investments in renewables and BESS rather than gas-fired power generation.

An important assumption is that treated and control regions are comparable in the absence of treatment. A common way to test this assumption is to explore trends before treatment. In Figure 9, we therefore examine dynamic treatment effects. We report effects for all plants in Subfigure (a), for renewable plants and BESS in Subfigure (b), and for gas-fired power plants in Subfigure (c). Because the dynamic version of the quarterly regressions is more susceptible to noise, we estimate these regressions using yearly data. Other than that, the specification follows Equation 1 with the single treatment indicator being replaced by leads and lags around the treatment year 2022; this year serves as the benchmark to capture dynamic treatment effects.

We find no evidence of pre-trends. In all pre-2022 years, the treatment effect is close to zero, except for 2018 where it is marginally negative. For 2023, we find no clear treatment effect yet, but in 2024 and 2025, there is a strong and positive effect, consistent with power plant investments taking some time to adjust to higher expected future demand after the ChatGPT shock.[25] Again this effect is driven by renewable plants and BESS, as we find no difference pre- and post-treatment for gas-fired power plants. These findings mitigate the concern that our results are driven by differences between treated and control regions that are unrelated to the treatment we exploit.

To understand which specific technology investments drive our results, we break down the effects for renewables with BESS into wind, solar, and BESS, and those for gas into grid-connected and BTM gas plants. For grid-connected gas, we consider both Simple-Cycle Natural Gas Turbines (SCGT) and Combined Cycle Gas Turbine (CCGT) plants. SCGT plants burn natural gas in a single turbine to generate electricity, while CCGT plants use both gas and steam turbines to generate electricity. The results in Table 3 show in Panel A for renewables that most of the treatment effect in data center regions comes from solar plants and BESS. There is no differential effect for wind plants. In Panel B, which reports results for gas plants, we observe that BTM plants become more frequent in data center regions after the release of ChatGPT, but this increase is moderated by a simultaneous decrease in grid-connected CCGT plants. Notably, a comparison of the effect sizes between BTM gas plants in Panel B and solar and BESS in Panel A shows much larger investment effects for renewables than for BTM gas plants.

[25] Additionally, as we measure the *announcement* of power plants, there might be a slight delay between the actual decision to invest in a power plant and its official announcement.

C. *Additional Tests and Robustness*

We perform two analyses to understand treatment intensity and to control for policy shocks at the state level. First, we replace the treatment dummy with treatment deciles to allow for a more continuous measure of treatment exposure. The results are reported in Table 4, Panel A. We focus on identifying effects for regions that locate in the 10^{th} and 9^{th} decile in terms of data center capacity as of September 2022; all other regions serve as the omitted benchmark category (for brevity, we report results only for $Capa_{r,t}^{0/1}$ and $Capa_{r,t}^{GW}$). If our previous interpretation is correct, we would expect the treatment effect to be concentrated in the 10^{th} decile and to be weaker in the 9^{th} decile. The results are consistent with this interpretation. Both for all plants and for renewable/BESS plants, the treatment effects are strongest for regions in the 10^{th} decile. The treatment coefficient estimate for the 9^{th} decile is smaller but statistically significant in three of the four specifications (and for all tests that consider renewables). These findings suggest that the treatment effect rises with treatment intensity, supporting our earlier conclusion that the AI boom led to greater investment in renewable energy and storage.

In Panel B, we include year-quarter $\times$ state fixed effects.[26] These high-dimensional fixed effects absorb both time-invariant differences across states and state-specific shocks that vary over time, substantially reducing concerns that omitted variables at the state level drive results. Identification now comes from variation across regions *within* the same state, while controlling for factors such as changes in state-level regulation, economic activity, or other developments that could affect investment in power generation. The results remain very similar to our baseline estimates: the release of ChatGPT increases power plant investment in treated regions relative to control regions, and this effect continues to be driven by renewable plants and BESS. The coefficient estimates for β are slightly smaller than in the baseline, but the differences are modest and all estimates remain statistically significant.

Robustness tests that address other potential concerns with our analysis are reported in Appendix C. In these specifications, we use $Capa_{r,t}^{0/1}$, which captures whether any new power plant investment occurred in region r at year-quarter t, as the dependent variable and report each robustness test as a separate row. For each test, we present the coefficient estimates for β for all plants, renewable plants with BESS, and gas-fired plants. The robustness checks can be grouped into three categories: those related to the treatment definition, timing-related tests, and alternative box sizes to identify regions.

[26] For each grid cell, we use the state in which most existing plants are located.

For the treatment-definition tests, we use the data center capacity as of 2022Q3 or the logarithm of this variable plus one as a continuous treatment measure. We also consider alternative binary treatment definitions, classifying a region as treated if its data center capacity in 2022Q3 falls within the top 20% or top 5%, or if the region has data center capacity above either zero or 100 MW in 2022Q3. Next, we address a potential forward-looking bias in our treatment classification. In the baseline specification, missing data center capacities in the 2022Q3 snapshot from the S&P 451 data are replaced with values from the 2025Q3 version. We re-estimate our model without this adjustment and alternatively use the number of data centers instead of total capacity. Across all these alternative definitions, results remain similar to the baseline.

For the timing-related tests, we restrict the sample to windows of ±10 or ±13 quarters around the ChatGPT shock (the baseline uses 2016 to 2025). We also exclude 2022 and 2023, when the effects of AI may not yet have been fully understood, and re-estimate specifications using yearly data. None of these changes impacts the results substantially.

For the box-size-related tests, we consider sizes of 50, 75, 200, and 300 km, as well as 50, 200, and 300 miles. For larger boxes, we replace $Capa_{r,t}^{0/1}$ with $Capa_{r,t}^{GW}$ since the indicator variable approaches one as the box size becomes very large. We find consistent results across all box sizes, although statistical significance declines for larger boxes, in line with our argument that power plants are typically built close to data centers.

D. Economic Mechanisms

We examine two channels that jointly may act as catalysts for increasing investments in renewables after the ChatGPT shock.

D.1. Power Purchase Agreements and Project Finance

As explained in Section II.B.1, the first mechanism relates to financial instruments used to facilitate the financing of investments in renewables. Specifically, one mechanism through which the AI boom may increase investment in renewable power plants and BESS is by expanding the availability of PPAs, which—as explained—are critical for project finance and are particularly important for renewables.

We test this mechanism in Table 5, Panel A. For this analysis, we need to turn from the region to the power plant level, to examine how finance-related plant characteristics change after the release of ChatGPT in treated and control plants. A power plant is defined as treated in these tests if a data center existed or was planned within a radius

of 10 km around it as of 2022Q3. The dependent variables of interest indicate whether a power plant uses a PPA (Columns 1 and 2) and whether it is project-financed (Columns 3 and 4). Because information on PPAs and project finance is likely incomplete for small power plants, we exclude plants with less than 10 MW of capacity from this analysis. We find that both the use of PPAs and the prevalence of project finance increased after the release of ChatGPT among treated plants relative to control plants. In terms of economic magnitude, the coefficient estimates suggest that the treatment increases the probability of a plant having a PPA by slightly less than 20 percentage points and the probability of a plant being project-financed by approximately 15 percentage points. This finding is consistent with the mechanism outlined above.

In Panel B, we test the implicit assumption underlying our proposed mechanism, namely that PPAs and project finance are more prevalent among renewable power plants and BESS. The key dependent variable is an indicator for whether a plant uses solar, wind, or BESS (zero for other technologies), and the independent variables of interest are indicators for whether the plant uses a PPA (Columns 1 and 2) or project finance (Columns 3 and 4). Consistent with our expectations and with the model of Darmouni et al. (2026), both PPAs and project finance are more common among plants using renewables and BESS. Plants that use PPAs are 25 percentage points more likely to use renewable technologies or BESS, while project-financed plants are 10 percentage points more likely to do so. These results provide further support for the proposed mechanism.

D.2. Climate Commitments by Data Center Operators

As explained in Section II.B.2, the second mechanism relates to the role of prior net-zero climate commitments by data center operators. To distinguish between data centers whose operators have made a climate pledge and those whose operators have not, we conduct this analysis at the data center level. Specifically, we examine how data center capacity affects nearby power plant investment before and after the release of ChatGPT, depending on whether the operator has made a climate commitment. Table 6 presents the results. Tn Panel A the dependent variable is $Capa_{d,t}^{0/1}$, which equals one if any new power plant became operational within a radius of 10 km in the same year as the data center's opening or within one year before or after. In Panel B, the dependent variable is constructed using only renewable power plants and BESS, and Panel C only for gas plants. Column 1 uses the full sample, Column 2 restricts the sample to data centers whose operators have a climate pledge, Column 3 restricts it to those without a climate

pledge, and Column 4 includes an interaction between $Post_t \times Log(DC\,capa_d)$ and an indicator for whether the operator has a climate pledge.

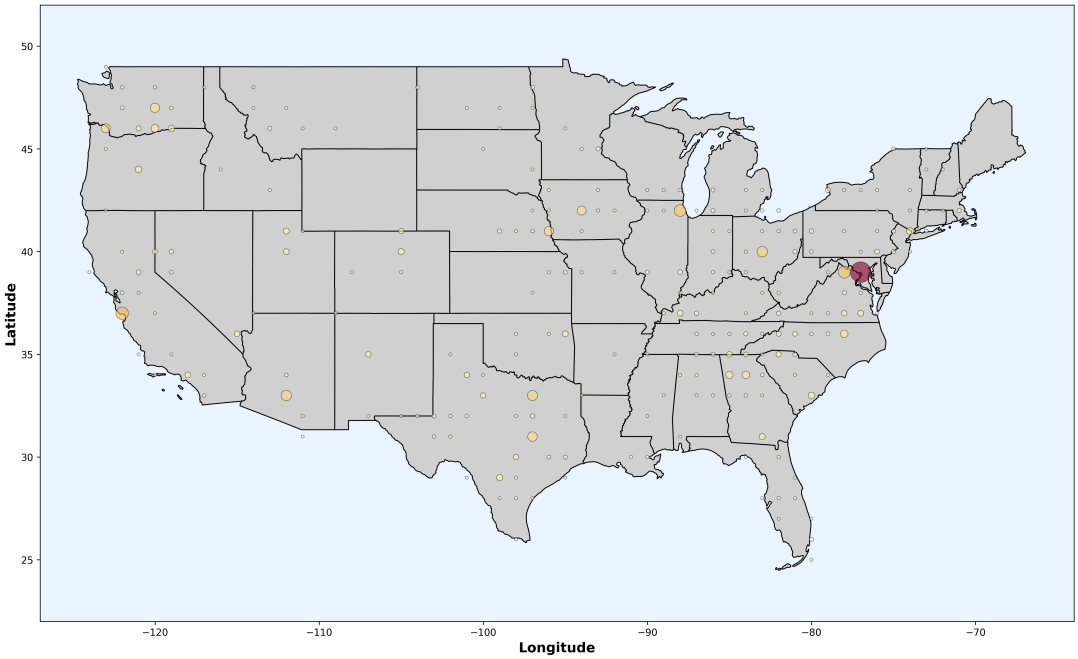
In Panel A, data center capacity has a positive effect on nearby overall power plant investment that is driven by the ChatGPT shock; this is in line with our earlier evidence. This result holds for operators both with and without a climate pledge. When we focus in Panel B specifically on renewable plants and BESS, the investment increase is concentrated around data centers whose operators have made a climate pledge. In Column 2, the coefficient estimate for $Post_t \times Log(DC\ capa_d)$ is large and positive among data centers with climate pledges (the effect is estimated with noise and marginally insignificant). To the contrary, in Column 3, the estimate for $Post_t \times Log(DC\ capa_d)$ is negative and statistically insignificant for firms without a climate pledge. Reflecting these sharp differences, we find in Column 4 that the triple interaction with $Climate\ pledge_d$ is statistically significant, indicating that renewable and storage investments are significantly more likely to occur around data centers whose operators made climate pledges.

When we turn to gas-fired power plants in Panel C, the opposite pattern emerges. The increase in newly added capacity is concentrated around data centers whose operators did not make a climate pledge, whereas the interaction term shows that significantly *less* gas capacity is added around data centers whose operators have made a climate pledge (compared to those whose operators have not). This contrast between cleaner technologies and gas-fired generation suggests that climate pledges matter not only for the scale of nearby power investment, but also for its composition. These results imply that climate pledges can have meaningful real effects on the technology mix of electricity supply and contribute to decarbonization, with PPAs playing a critical role for this effect.

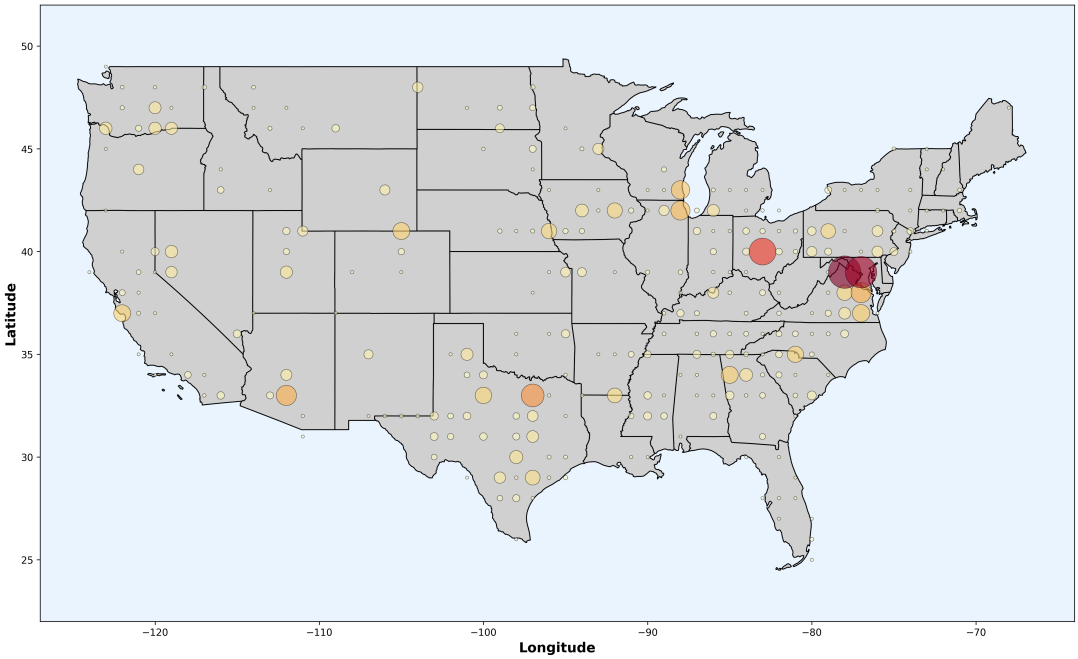
VI. Conclusions

This paper examines how, and through which economic channels, the sharp increase in electricity demand due to the AI revolution shapes investment in different types of power generation. Using the public release of ChatGPT in November 2022 as an exogenous shock, we show that planned power plant capacity increases significantly more in regions with data centers than in those without. Contrary to widespread belief, this increase is driven primarily by renewable power plants and battery storage rather than gas-fired generation. We provide evidence that data centers' use of PPAs—which facilitate project financing for renewable energy—and net-zero climate commitments by hyperscalers play a key role in this expansion. Our results shed light on how economies respond to shocks in

electricity demand, providing a useful laboratory for understanding the dynamics of large-scale electrification needed for the net-zero transition. A caveat is that data center-related renewable energy investments might crowd out or replace corresponding investments in other industries that are less deep-pocketed.



(a) U.S. Data Center Power Capacity (as of 2022Q3)



(b) U.S. Data Center Power Capacity (as of 2025Q3)

Figure 1: This figure shows all operating and planned U.S. data centers in the S&P 451 Research's Datacenter KnowledgeBase by capacity and location. Subfigure (a) contains data as of 2022Q3 (before the ChatGPT shock) and Subfigure (b) as of 2025Q3 (after the ChatGPT shock). Each circle aggregates data center capacity at a latitude/longitude pair (rounded to full numbers; e.g., 39°N & 79°W), and its size and color indicate the power capacity of all planned and operating data centers at the latitude/longitude pair in MW. The larger the circle, the darker the shade of red.

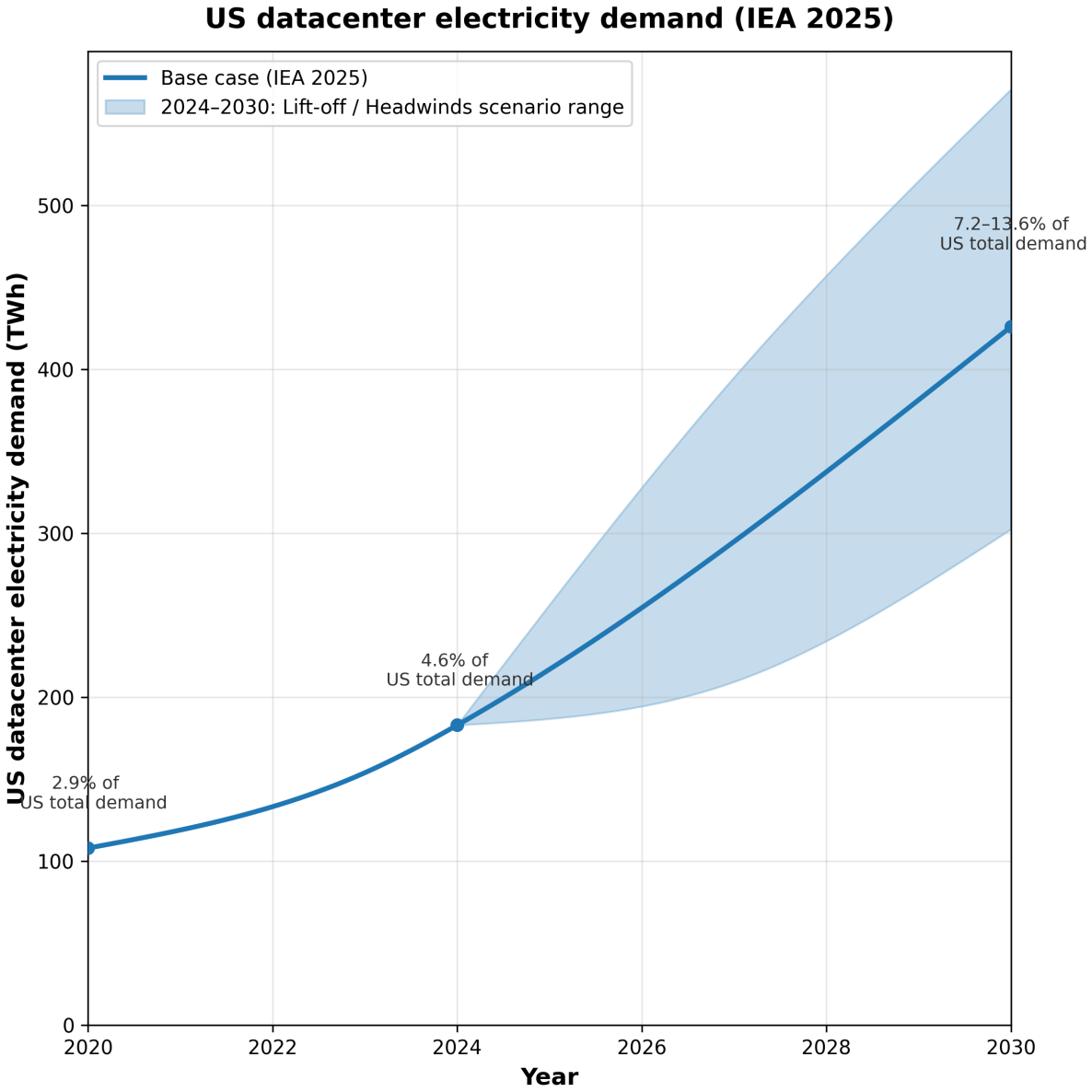


Figure 2: This figure shows U.S. data center electricity demand based on historical and forecasted data from International Energy Agency (2025). The share of data center demand relative to total U.S. electricity demand is based on electricity sales figures from the U.S. Energy Information Administration's Annual Energy Outlook 2025. The data center power demand has increased from 108 TWh in 2020 to 183 TWh in 2024 and is estimated to rise to 426 TWh in IEA's base case. This leads to an increase in the share of U.S. power demand from 2.9% in 2020 to between 7.2 and 13.6% in 2030.

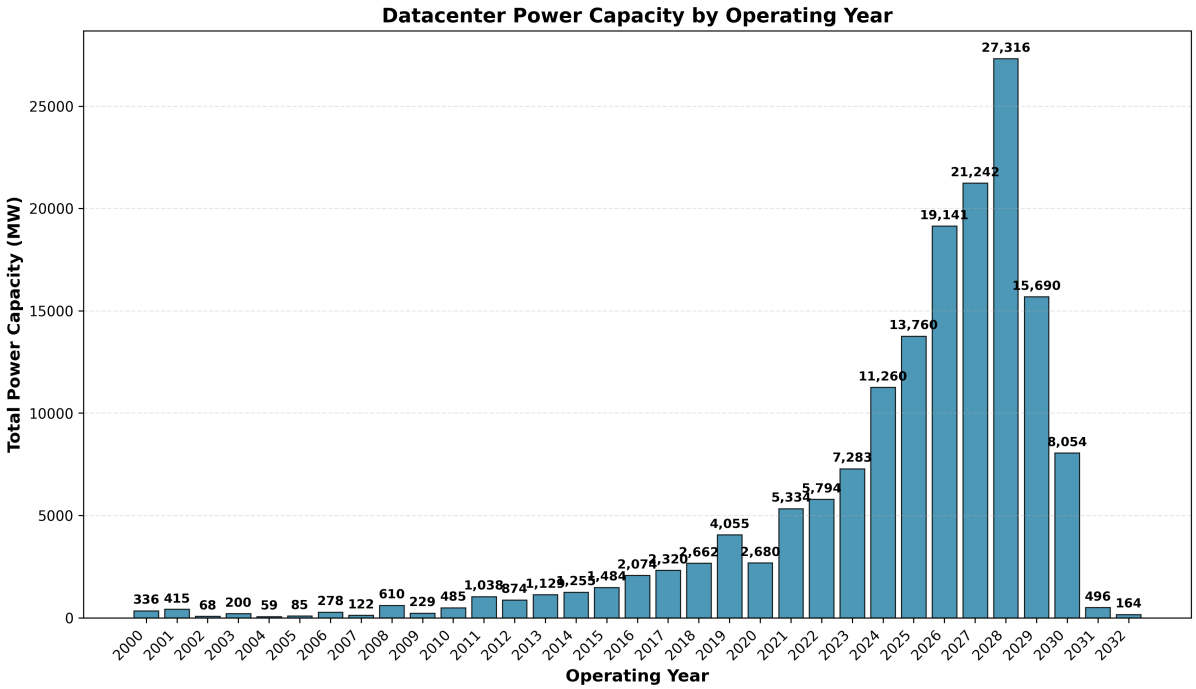


Figure 3: This figure shows the total power capacity of U.S. data centers (in MW) by operating year according to the S&P 451 Research's Datacenter KnowledgeBase. The data is as of 2025Q3. Data center power capacity after 2025Q3 is based on the planned start date of operation.

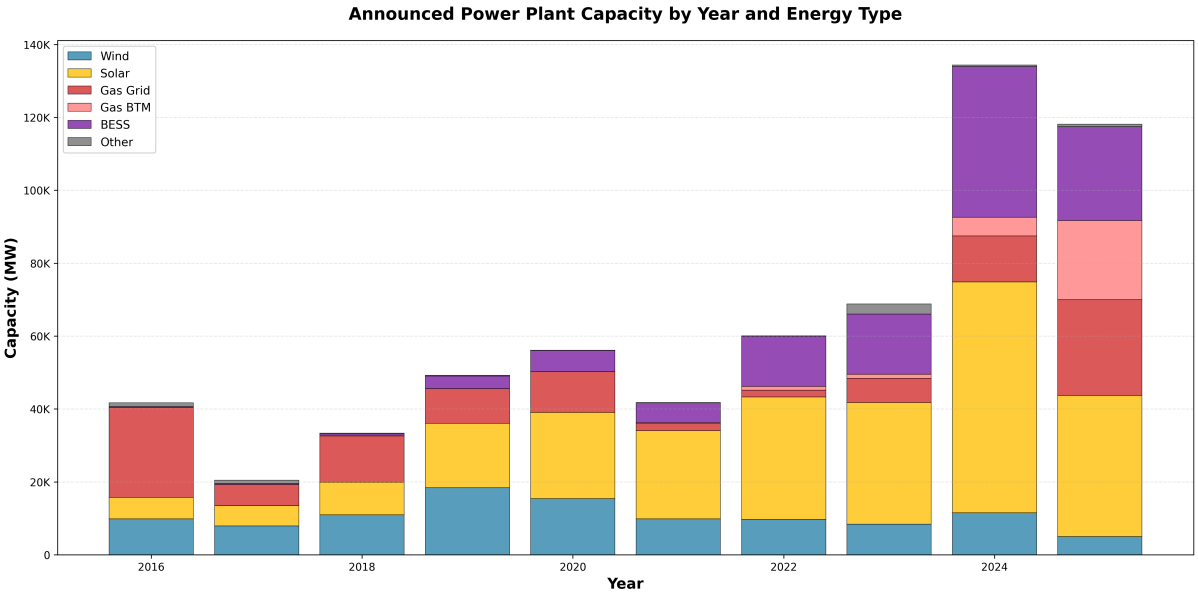


Figure 4: This figure shows the capacity of all power plants included in our sample (in MW) by announcement year and power plant type. BESS stands for Battery Energy Storage Systems and BTM for behind-the-meter energy production.

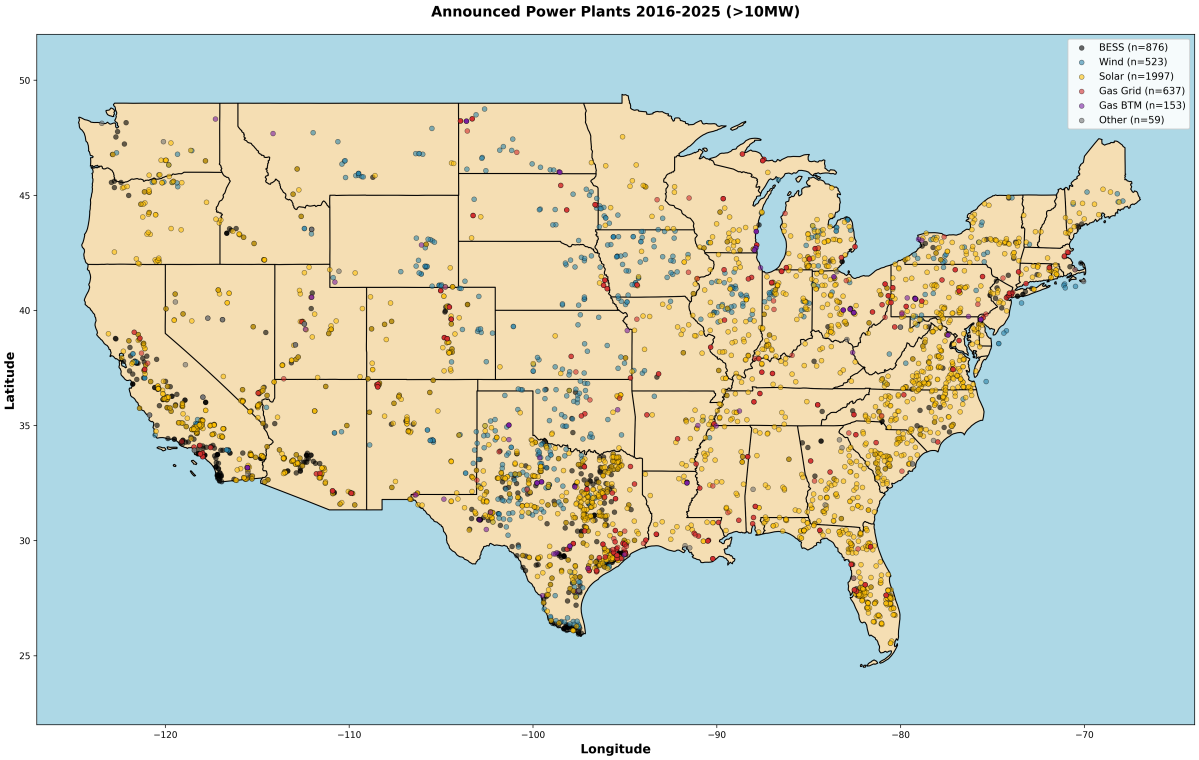


Figure 5: This figure shows all power plant units with more than 10 MW capacity that were announced during our sample period from 2016 to 2025. The figure depicts these plants by type and location. BESS stands for Battery Energy Storage Systems and BTM for behind-the-meter energy production.

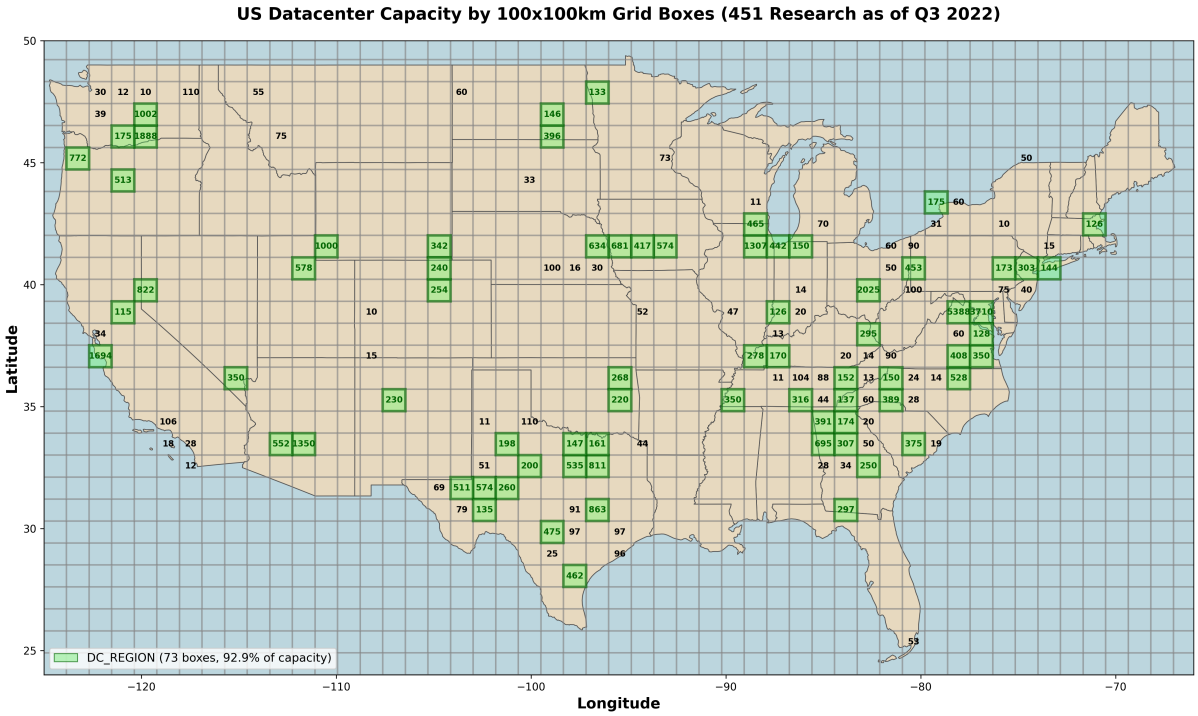
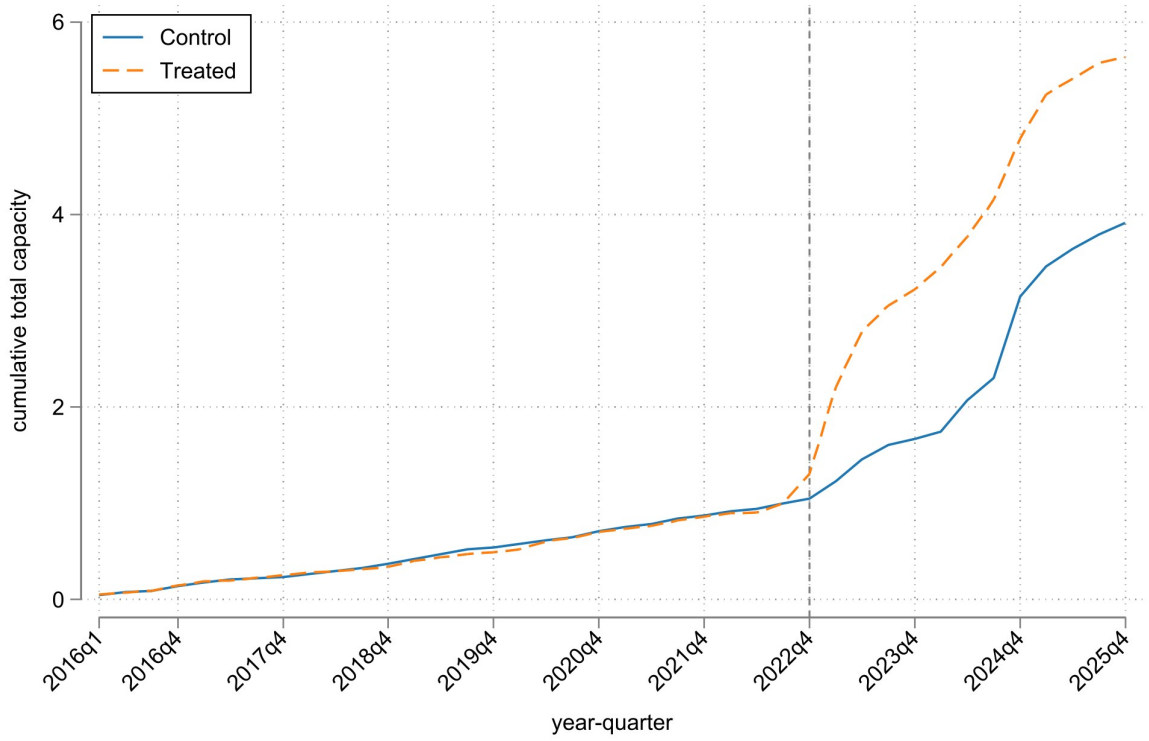
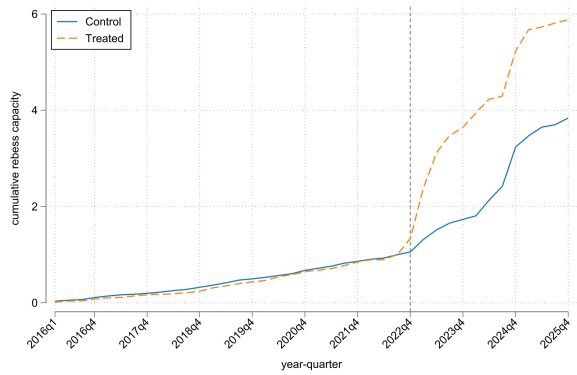


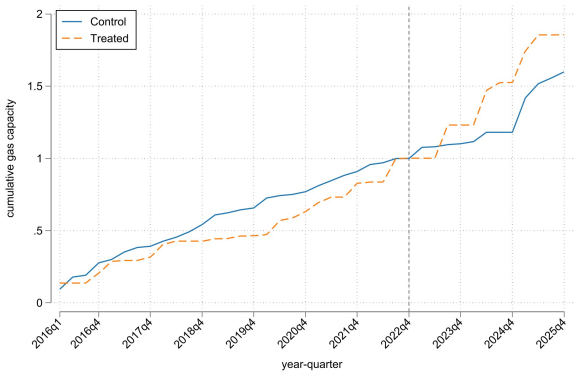
Figure 6: This figure shows the capacity of all operating and planned U.S. data centers as of 2022Q3. The U.S. is split up into approximately 100 × 100km boxes. New power-generating capacity is typically built close to data center locations. Green boxes indicate data center regions, defined as the top 10% of location boxes by data center capacity (across the U.S. mainland). The numbers in each cell indicate the respective data center capacity (in MW). The 73 data center region boxes in the top 10% contain 92.9% of the total U.S. data center capacity as of 2022Q3.



(a) All Power Plants

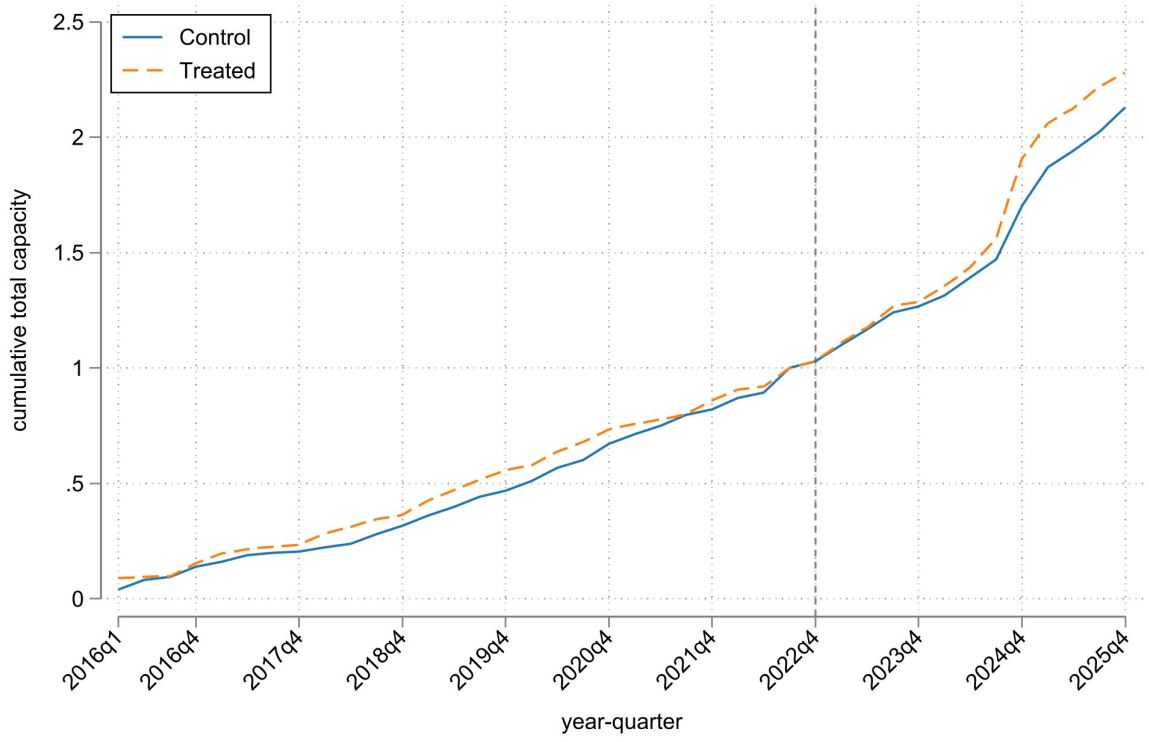


(b) Renewable Plants and BESS

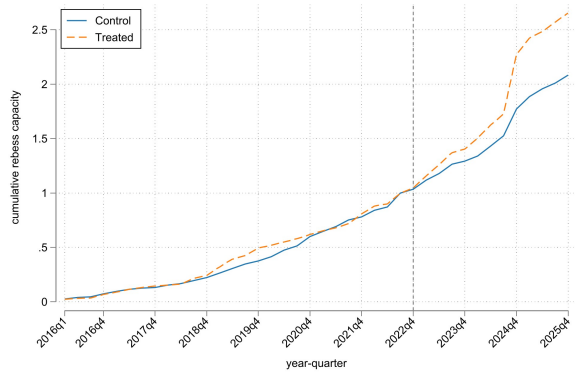


(c) Gas-Fired power plants

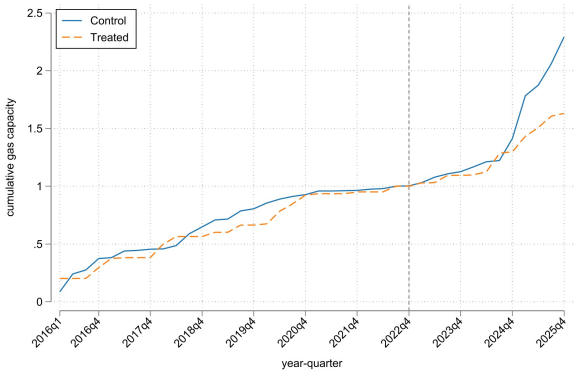
Figure 7: This figure shows trends of the average announced cumulative power plant capacity in data center regions (treated) vs. non-data center regions (control). The announced capacity is first summed for each region. After that, the region-specific cumulative capacity in a year-quarter is averaged and normalized at 100% in 2022Q3 for treated and control regions, respectively. Subfigure (a) considers all power plants, Subfigure (b) only renewable plants and battery energy storage systems (BESS), and Subfigure (c) only gas-fired power plants.



(a) All Power Plants

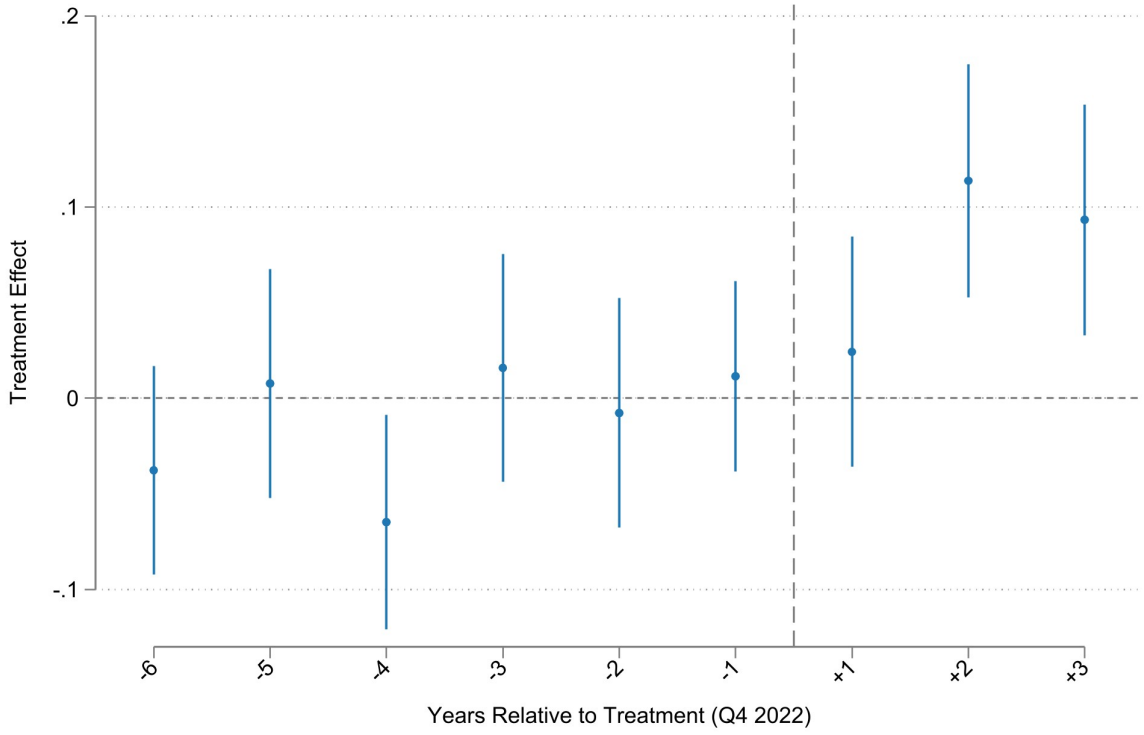


(b) Renewable Plants and BESS

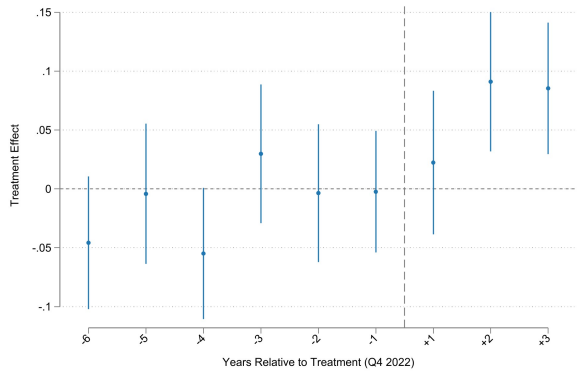


(c) Gas-Fired Power Plants

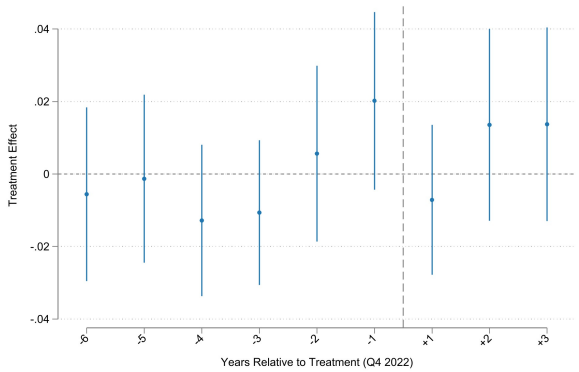
Figure 8: This figure shows trends of the aggregate announced cumulative power plant capacity in data center regions (treated) vs. non-data center regions (control). The announced capacity is summed across all treated and control regions in a year-quarter and normalized at 100% in 2022Q3. Subfigure (a) considers all power plants, Subfigure (b) only renewable plants and battery energy storage systems (BESS), and Subfigure (c) only gas-fired power plants.



(a) All Power Plants



(b) Renewable Plants and BESS



(c) Gas-Fired Power Plants

Figure 9: This figure shows dynamic yearly treatment effects for the region-level analysis. The underlying regression specification follows Column 1 of Table 2. For this visual figure, we estimate the model on the year-level to reduce noise. Subfigure (a) considers all power plants, Subfigure (b) only renewable plants and battery energy storage systems (BESS), and Subfigure (c) only gas-fired power plants.

Table 1: Descriptive Statistics

Variable	Obs.	Mean	p25	p50	p75	SD
Panel A: Region-Level Variables						
Total Capacity						
$\text{Capa}_{r,t}^{0/1}$	29,253	0.13	0.00	0.00	0.00	0.34
$\text{Capa}_{r,t}^{GW}$	29,253	0.02	0.00	0.00	0.00	0.12
$\text{Log}(\text{Capa}_{r,t})$	3,902	3.57	1.69	3.91	5.29	1.98
$\text{Log}(1+\text{Capa}_{r,t})$	29,253	0.49	0.00	0.00	0.00	1.42
RE/BESS Capacity						
$\text{Capa}_{r,t}^{0/1}$	29,253	0.12	0.00	0.00	0.00	0.32
$\text{Capa}_{r,t}^{GW}$	29,253	0.02	0.00	0.00	0.00	0.09
$\text{Log}(\text{Capa}_{r,t})$	3,509	3.50	1.61	3.91	5.17	1.93
$\text{Log}(1+\text{Capa}_{r,t})$	29,253	0.43	0.00	0.00	0.00	1.33
Gas Capacity						
$\text{Capa}_{r,t}^{0/1}$	29,253	0.01	0.00	0.00	0.00	0.11
$\text{Capa}_{r,t}^{GW}$	29,253	0.00	0.00	0.00	0.00	0.08
$\text{Log}(\text{Capa}_{r,t})$	365	4.51	2.77	4.88	6.38	2.18
$\text{Log}(1+\text{Capa}_{r,t})$	29,253	0.06	0.00	0.00	0.00	0.56
Other Variables						
Treat_r	29,253	0.10	0.00	0.00	0.00	0.30
Post_t	29,253	0.33	0.00	0.00	1.00	0.47
Plant $\text{capa}_{r,t}$	29,253	2.18	0.25	1.12	2.96	2.90
$\text{Log}(\text{Plant capa}_{r,t})$	29,253	6.60	5.52	7.02	7.99	1.88
Share $\text{RE/BESS}_{r,t}$	29,253	0.12	0.00	0.05	0.26	0.13
Panel B: Plant-Level Variables						
$\text{PPA}_p^{0/1}$	4,179	0.39	0.00	0.00	1.00	0.49
$\text{Project finance}_p^{0/1}$	4,179	0.35	0.00	0.00	1.00	0.48
Treat_p	4,283	0.08	0.00	0.00	0.00	0.28
Post_t	4,283	0.49	0.00	0.00	1.00	0.50
$\text{RE/BESS}_p^{0/1}$	4,283	0.80	1.00	1.00	1.00	0.40
Plant capa_p	4,283	142.95	50.00	100.00	200.00	162.26
$\text{Log}(\text{Plant capa}_p)$	4,283	4.47	3.91	4.61	5.30	1.05

Panel C: Data-Center-Level Variables						
Total Capacity						
$\text{Capa}_{d,t}^{0/1}$	2,126	0.26	0.00	0.00	1.00	0.44
RE/BESS Capacity						
$\text{Capa}_{d,t}^{0/1}$	2,126	0.17	0.00	0.00	0.00	0.38
Gas Capacity						
$\text{Capa}_{d,t}^{0/1}$	2,126	0.12	0.00	0.00	0.00	0.33
Other Variables						
DC capa_d	2,067	75.28	29.91	50.00	78.00	122.82
Log(DC capa_d)	2,067	3.91	3.40	3.91	4.36	0.82
Post_t	2,126	0.69	0.00	1.00	1.00	0.46
Climate pledge_d	2,040	0.66	0.00	1.00	1.00	0.47

This table presents descriptive statistics for region-level variables (Panel A), plant-level variables (Panel B), and data-center-level variables (Panel C). The sample period is from 2016 to 2025. A detailed description of all variables can be found in Appendix A.

Table 2: Power Plant Investments and Data Centers: Main Results

Panel A: All Power Plants

Dependent var.:	(1) $\text{Capa}_{r,t}^{0/1}$	(2) $\text{Capa}_{r,t}^{GW}$	(3) $\text{Log}(\text{Capa}_{r,t})$	(4) $\text{Log}(1+\text{Capa}_{r,t})$
$\text{Post}_t \times \text{Treat}_r$	0.082***	0.030***	0.14	0.45***
	(3.45)	(2.61)	(0.76)	(3.82)
$\text{Log}(\text{Plant capa}_{r,t})$	-0.027***	-0.0077***	-0.34***	-0.13***
	(-3.69)	(-2.85)	(-2.91)	(-3.97)
$\text{Share RE/BESS}_{r,t}$	0.23***	0.12***	1.61*	1.37***
	(2.68)	(3.21)	(1.69)	(3.31)
Observations	29,252	29,252	3,790	29,252
Regions	750	750	499	750
Adj. R^2	0.19	0.077	0.36	0.16
Region FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes

Panel B: Renewable Power Plants and BESS

Dependent var.:	(1) $\text{Capa}_{r,t}^{0/1}$	(2) $\text{Capa}_{r,t}^{GW}$	(3) $\text{Log}(\text{Capa}_{r,t})$	(4) $\text{Log}(1+\text{Capa}_{r,t})$
$\text{Post}_t \times \text{Treat}_r$	0.075***	0.031***	0.29*	0.43***
	(3.19)	(3.32)	(1.71)	(3.78)
$\text{Log}(\text{Plant capa}_{r,t})$	-0.024***	-0.0033	-0.23**	-0.10***
	(-3.51)	(-1.47)	(-2.03)	(-3.52)
$\text{Share RE/BESS}_{r,t}$	0.20**	0.070**	0.28	1.11***
	(2.45)	(2.32)	(0.30)	(2.86)
Observations	29,252	29,252	3,391	29,252
Regions	750	750	461	750
Adj. R^2	0.19	0.096	0.42	0.16
Region FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes

continued on next page

Table 2 continued

Panel C: Gas-Fired Power Plants

Dependent var.:	(1) $\text{Capa}_{r,t}^{0/1}$	(2) $\text{Capa}_{r,t}^{GW}$	(3) $\text{Log}(\text{Capa}_{r,t})$	(4) $\text{Log}(1+\text{Capa}_{r,t})$
$\text{Post}_t \times \text{Treat}_r$	0.0052	-0.0017	-0.29	0.032
	(0.56)	(-0.31)	(-0.39)	(0.66)
$\text{Log}(\text{Plant capa}_{r,t})$	-0.0045**	-0.0043***	-1.14*	-0.027**
	(-2.10)	(-2.74)	(-1.88)	(-2.33)
$\text{Share RE/BESS}_{r,t}$	0.059**	0.048**	5.63	0.38***
	(2.30)	(2.27)	(1.21)	(2.59)
Observations	29,252	29,252	241	29,252
Regions	750	750	80	750
Adj. R^2	0.045	0.021	0.46	0.036
Region FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes

This table shows how power plant investments changed around the introduction of ChatGPT in November 2022. The analysis is conducted at the region level. The sample period is from 2016 to 2025. The dependent variable in Column 1 is $Capa_{r,t}^{0/1}$, which equals one if any new power plant is announced in year-quarter t and region r. In Column 2, the dependent variable is $Capa_{r,t}^{GW}$, which is the capacity of all newly announced power plants in year-quarter t and region r (in GW). The dependent variables in Columns 3 and 4 are $(Capa_{r,t})$ and $(1 + Capa_{r,t})$, respectively. We use capacity in MW for the logarithmic specifications. We only consider renewable power plants and battery energy storage systems when constructing the dependent variables in Panel B, and gas-fired power plants in Panel C. $Treat_r$ is an indicator variable equal to one for regions in the top 10% of data center capacity as of September 30, 2022 (before the release of ChatGPT). $Post_t$ is an indicator variable equal to one for observations from 2022Q4 onward. t-statistics based on robust standard errors clustered by regions are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 3: Power Plant Investments and Data Centers: Technology-Specific Results

Panel A: Renewable Power Plants and BESS

	(1)	(2)	(3)	(4)	(5)	(6)
Technology	Solar		Wind		BESS	
Dependent var.:	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$
$\text{Post}_t \times \text{Treat}_r$	0.042*	0.017***	0.0017	-0.00024	0.053***	0.014***
	(1.75)	(3.21)	(0.32)	(-0.15)	(3.58)	(2.85)
$\text{Log(Plant capa}_{r,t})$	-0.013**	-0.00044	-0.016***	-0.0039**	0.0065*	0.0011
	(-2.22)	(-0.34)	(-4.86)	(-2.07)	(1.78)	(1.52)
$\text{Share RE/BESS}_{r,t}$	0.18**	0.038**	0.015	0.0064	0.091*	0.025*
	(2.47)	(2.25)	(0.54)	(0.48)	(1.76)	(1.93)
Observations	29,252	29,252	29,252	29,252	29,252	29,252
Adj. R^2	0.18	0.087	0.058	0.027	0.13	0.089
Region FE	yes	yes	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes	yes	yes

Panel B: Gas-Fired Power Plants

	(1)	(2)	(3)	(4)	(5)	(6)
Technology	Grid SCGT		Grid CCGT		BTM	
Dependent var.:	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$
$\text{Post}_t \times \text{Treat}_r$	-0.0083	-0.0028	-0.0095***	-0.0059***	0.016***	0.0069*
	(-1.16)	(-1.10)	(-3.14)	(-2.74)	(2.72)	(1.86)
$\text{Log(Plant capa}_{r,t})$	-0.0055***	-0.0026***	-0.0023**	-0.0013*	0.00083	-0.00036
	(-2.78)	(-3.25)	(-1.97)	(-1.80)	(1.02)	(-0.81)
$\text{Share RE/BESS}_{r,t}$	0.045**	0.020**	0.026*	0.016	0.016	0.012
	(2.12)	(2.18)	(1.91)	(1.53)	(1.14)	(1.28)
Observations	29,252	29,252	29,252	29,252	29,252	29,252
Adj. R^2	0.040	0.015	0.014	0.011	0.020	0.0052
Region FE	yes	yes	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes	yes	yes

continued on next page

Table 3 continued

This table shows how investments in different power plant technologies changed around the introduction of ChatGPT in November 2022. The analysis is conducted at the region level. The sample period is from 2016 to 2025. The dependent variable in uneven columns is $Capa_{r,t}^{0/1}$, which equals one if any new power plant is announced in year-quarter t and region r. The dependent variable in even columns is $Capa_{r,t}^{GW}$, which equals the capacity of all newly announced power plants in year-quarter t and region r (in GW). Only power plants using the technology specified in each column are considered when constructing the dependent variables. BESS stands for Battery Energy Storage Systems. SCGT stands for Simple-Cycle Natural Gas Turbines plants, which are power plants that burn natural gas in a single turbine to generate electricity. CCGT stands for Combined Cycle Gas Turbine plants, which using both gas and steam turbines to generate electricity. BTM stands for behind-the-meter gas plants, which are not connected to the grid. $Treat_r$ is an indicator variable equal to one for regions in the top 10% of data center capacity as of September 30, 2022 (before the release of ChatGPT). $Post_t$ is an indicator variable equal to one for observations from 2022Q4 onward. t-statistics based on robust standard errors clustered by regions are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 4: Power Plant Investments and Data Centers: Additional Analyses

Panel A: Treatment Deciles

	(1)	(2)	(3)	(4)	(5)	(6)
Technology	All		RE/BESS		Gas	
Dependent var.:	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$
$\text{Post}_t \times \text{Treat}_r^{\text{9th decile}}$	0.034	0.018*	0.046*	0.014*	-0.0052	0.0042
	(1.40)	(1.87)	(1.78)	(1.71)	(-0.82)	(0.91)
$\text{Post}_t \times \text{Treat}_r^{\text{10th decile}}$	0.085***	0.031***	0.080***	0.032***	0.0047	-0.0013
	(3.60)	(2.78)	(3.40)	(3.48)	(0.50)	(-0.23)
$\text{Log(Plant capa}_{r,t})$	-0.026***	-0.0073***	-0.023***	-0.0030	-0.0046**	-0.0042***
	(-3.66)	(-2.73)	(-3.47)	(-1.36)	(-2.14)	(-2.68)
$\text{Share RE/BESS}_{r,t}$	0.22***	0.12***	0.20**	0.067**	0.060**	0.048**
	(2.65)	(3.18)	(2.40)	(2.27)	(2.33)	(2.23)
Observations	29,252	29,252	29,252	29,252	29,252	29,252
Adj R^2	0.19	0.078	0.19	0.096	0.045	0.021
Region FE	yes	yes	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes	yes	yes

Panel B: High-Dimensional Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Technology	All		RE/BESS		Gas	
Dependent var.:	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$	$\text{Capa}_{r,t}^{0/1}$	$\text{Capa}_{r,t}^{GW}$
$\text{Post}_t \times \text{Treat}_r$	0.063***	0.021*	0.051***	0.024***	0.0062	-0.0033
	(3.08)	(1.85)	(2.63)	(2.60)	(0.65)	(-0.56)
$\text{Log(Plant capa}_{r,t})$	-0.030***	-0.010***	-0.026***	-0.0056**	-0.0067***	-0.0047***
	(-4.89)	(-3.59)	(-4.37)	(-2.36)	(-3.07)	(-3.30)
$\text{Share RE/BESS}_{r,t}$	0.13*	0.12***	0.092	0.066**	0.064***	0.049**
	(1.68)	(3.19)	(1.23)	(2.20)	(2.65)	(2.45)
Observations	29,172	29,172	29,172	29,172	29,172	29,172
Adj. R^2	0.24	0.096	0.25	0.11	0.073	0.036
Region FE	yes	yes	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes	yes	yes
Year-Quarter FE × State FE	yes	yes	yes	yes	yes	yes

continued on next page

Table 4 continued

This table shows additional specifications that estimate how power plant investments changed around the introduction of ChatGPT. The analysis is conducted at the region level. The sample period is from 2016 to 2025. In Panel A, we include dummies for the 9^{th} and 10^{th} decile of the data center capacity in region r as of September 30, 2022. Panel B shows specifications with high-dimensional state times year-quarter fixed effects. The dependent variable in uneven columns is $Capa^{0/1}_{r,t}$, which equals one if any new power plant is announced in year-quarter t and region r. The dependent variable in even columns is $Capa^{GW}_{r,t}$, which equals the capacity of all newly announced power plants in year-quarter t and region r (in GW). Only power plants using the technology specified in each column are considered when constructing the dependent variables. $Treat_r$ is an indicator variable equal to one for regions in the top 10% of data center capacity as of September 30, 2022 (before the release of ChatGPT). $Post_t$ is an indicator variable equal to one for observations from 2022Q4 onward. t-statistics based on robust standard errors clustered by regions are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 5: Economic Mechanism: Project Finance and Power Purchase Agreements

Panel A: Data Centers, PPAs, and Project Finance				
	(1)	(2)	(3)	(4)
Dependent var.:	$PPA_p^{0/1}$		$Project\ finance_p^{0/1}$	
$Post_t \times Treat_p$	0.18**	0.17**	0.15**	0.12**
	(2.40)	(2.31)	(2.49)	(2.17)
$Treat_p$	-0.23***	-0.22***	-0.14***	-0.095***
	(-4.97)	(-4.67)	(-3.96)	(-2.77)
$Log(Plant\ capa_p)$		0.018		0.083***
		(1.02)		(7.03)
Observations	4,179	4,179	4,179	4,179
Adj. R^2	0.19	0.19	0.091	0.12
State FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes
Panel B: PPAs, Project Finance, and Power Plant Technology				
	(1)	(2)	(3)	(4)
Dependent var.:			$RE/BESS_p^{0/1}$	
$PPA_p^{0/1}$	0.25***	0.25***		
	(10.0)	(10.00)		
$Project\ finance_p^{0/1}$			0.10***	0.10***
			(3.40)	(3.06)
$Log(Plant\ capa_p)$		-0.00063		-0.0032
		(-0.031)		(-0.12)
Observations	4,179	4,179	4,179	4,179
Adj. R^2	0.26	0.26	0.19	0.19
State FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes

continued on next page

Table 5 continued

This table examines the role of hedging through power purchase agreements (PPAs) and financing choices in explaining how power plant investments changed around the introduction of ChatGPT. The analysis is conducted at the power plant level. The sample period is from 2016 to 2025. The dependent variables in Panel A are variables that indicate whether a power plant uses a PPA and whether it is project financed. In Panel B, the dependent variable is an indicator of whether the plant's technology is wind, solar, or BESS. $Treat_p$ is an indicator variable that equals one for power plant p if any data center existed or was planned on September 30, 2022 (before the release of ChatGPT) within a radius of 10 km. $Post_t$ is an indicator variable equal to one for observations from 2022Q4 onward. Regions are equal to U.S. states for this analysis. Only power plants with a capacity of at least 10 MW are included in the analysis. t-statistics based on robust standard errors clustered by U.S. states are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 6: Economic Mechanism: Climate Commitments of Data Center Operators

Panel A: All Power Plants				
	(1)	(2)	(3)	(4)
Dependent var:			$\text{Capa}_{d,t}^{0/1}$	
Sample	All	Pledge	No pledge	All
$\text{Log(DC capa}_d\text{)}$	-0.030	-0.037	-0.048	-0.015
	(-1.68)	(-1.13)	(-1.43)	(-0.60)
$\text{Post}_t \times \text{Log(DC capa}_d\text{)}$	0.095***	0.098**	0.13**	0.093*
	(2.74)	(2.30)	(2.16)	(1.94)
Climate pledge_d				0.065
				(0.48)
$\text{Post}_t \times \text{Climate pledge}_d \times \text{Log(DC capa}_d\text{)}$				-0.012
				(-0.20)
Observations	2,054	1,308	662	1,986
Adj. R^2	0.22	0.24	0.31	0.23
State FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes
Panel B: Renewable Power Plants and BESS				
	(1)	(2)	(3)	(4)
Dependent var:			$\text{Capa}_{d,t}^{0/1}$	
Sample	All	Pledge	No pledge	All
$\text{Log(DC capa}_d\text{)}$	-0.015	-0.022	-0.0024	0.011
	(-1.15)	(-1.00)	(-0.085)	(0.72)
$\text{Post}_t \times \text{Log(DC capa}_d\text{)}$	0.0066	0.045	-0.027	-0.051*
	(0.40)	(1.68)	(-0.78)	(-1.94)
Climate pledge_d				0.12
				(1.26)
$\text{Post}_t \times \text{Climate pledge}_d \times \text{Log(DC capa}_d\text{)}$				0.100**
				(2.48)
Observations	2,054	1,308	662	1,986
Adj. R^2	0.16	0.18	0.20	0.17
State FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes

continued on next page

Table 6 continued

Panel C: Gas-Fired Power Plants

	(1)	(2)	(3)	(4)
Dependent var:			$\text{Capa}_{d,t}^{0/1}$	
Sample	All	Pledge	No pledge	All
Log(DC capa_d)	-0.0099	0.0052	-0.062**	-0.026
	(-0.79)	(0.26)	(-2.37)	(-1.40)
$\text{Post}_t \times$ Log(DC capa_d)	0.085***	0.052	0.14***	0.13***
	(2.73)	(1.62)	(3.70)	(3.36)
Climate pledge_d				-0.15
				(-1.59)
$\text{Post}_t \times$ Climate $\text{pledge}_d \times$ Log(DC capa_d)				-0.099**
				(-2.20)
Observations	2,054	1,308	662	1,986
Adj. R^2	0.32	0.29	0.43	0.33
State FE	yes	yes	yes	yes
Year-Quarter FE	yes	yes	yes	yes

This table shows power plant investments around newly opened data centers. The analysis is conducted at the data center level. The sample period is from 2016 to 2025. The dependent variable in Panel A is $Capa_{d,t}^{0/1}$, which equals one if any new power plant became operational within a radius of 10 km around data center d in the same year or one year before/after the data center's opening. We only consider renewable power plants and battery energy storage systems when constructing the dependent variable in Panel B, and gas-fired power plants in Panel C. $Post_t$ is an indicator variable equal to one for observations from 2022Q4 onward. Climate pledge is an indicator variable that equals one if the operator of data center d has a climate pledge and zero otherwise. Only data centers with a capacity of at least 10 MW are included in the analysis. t-statistics based on robust standard errors clustered by U.S. states are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Appendix

Appendix A: Definition of Variables

Variable	Description
Panel A: Region-Level Variables	
Total Capacity	
$\text{Capa}_{r,t}^{0/1}$	Indicator variable that equals one if any new power plant is announced in year-quarter t and region r. Source: U.S. EIA Form 860M.
$\text{Capa}_{r,t}^{GW}$	Capacity of all newly announced power plants in year-quarter t and region r (in GW). Source: U.S. EIA Form 860M.
RE/BESS Capacity	
$\text{Capa}_{r,t}^{0/1}$	Indicator variable that equals one if any new renewable power plant or battery energy storage system is announced in year-quarter t and region r. Source: U.S. EIA Form 860M.
$\text{Capa}_{r,t}^{GW}$	Capacity of all newly announced renewable power plants and battery energy storage systems in year-quarter t and region r (in GW). Source: U.S. EIA Form 860M.
Gas Capacity	
$\text{Capa}_{r,t}^{0/1}$	Indicator variable that equals one if any new gas-fired power plant is announced in year-quarter t and region r. Source: U.S. EIA Form 860M.
$\text{Capa}_{r,t}^{GW}$	Capacity of all newly announced gas-fired power plants in year-quarter t and region r (in GW). Source: U.S. EIA Form 860M.
Other Variables	
Treat_r	Indicator variable equal to one for regions in the top 10% of data center capacity as of September 30, 2022. Source: 451 Research Datacenter KnowledgeBase.
Post_t	Indicator variable that equals one for observations from 2022 Q4 onward.
Plant $\text{capa}_{r,t}$	Capacity of all existing power plants in year-quarter t and region r (in MW). Source: U.S. EIA Form 860M.
Share $\text{RE/BESS}_{r,t}$	Capacity of all renewable power plants and battery energy storage systems in year-quarter t and region r scaled by the total existing capacity. Source: U.S. EIA Form 860M.
Panel B: Plant-Level Variables	
$\text{PPA}_p^{0/1}$	Indicator variable that equals one if power plant p uses a PPA. Source: CapitalIQ Pro
Project $\text{finance}_p^{0/1}$	Indicator variable that equals one if power plant p is project financed. Source: LSEG Refinitiv
Treat_p	Indicator variable that equals one for power plant p if any data center existed or was planned on September 30, 2022 within a radius of 10 km. Source: 451 Research Datacenter KnowledgeBase.

continued on next page

Appendix A continued

Variable	Description
$\text{RE/BESS}_p^{0/1}$	Indicator variable that equals one for power plant p if its technology is wind, solar, or BESS. Source: U.S. EIA Form 860M.
Plant capa_p	Capacity of power plant p. Source: U.S. EIA Form 860M.
Panel C: Data-Center-Level Variables	
Total Capacity	
$\text{Capa}_{d,t}^{0/1}$	Indicator variable that equals one if any new power plant became operational within a radius of 10 km around data center d in the same year or one year before/after the data center's opening. Source: U.S. EIA Form 860M.
RE/BESS Capacity	
$\text{Capa}_{d,t}^{0/1}$	Indicator variable that equals one if any new renewable power plant or battery energy storage system became operational within a radius of 10 km around data center d at the same year or one year before/after the data center's opening. Source: U.S. EIA Form 860M.
Gas Capacity	
$\text{Capa}_{d,t}^{0/1}$	Indicator variable that equals one if any new gas-fired power plant became operational within a radius of 10 km around data center d at the same year or one year before/after the data center's opening. Source: U.S. EIA Form 860M.
Other Variables	
DC capa_d	Total utility power capacity of data center d in MW. Source: 451 Research Datacenter KnowledgeBase.
Climate pledge_d	Indicator variable that equals one if the operator of data center d has publicly announced a carbon net-zero target year. Source: Press releases and statements on websites of data center operators.

This table provides definitions of the main variables.

Appendix B: Regional Clustering of Data Centers

Dependent var.:	(1) DC capa$_r^{\geq 2023Q1,0/1}$	(2) DC capa$_r^{\geq 2023Q1}$	(3) Log(1+DC capa$_r^{\geq 2023Q1}$)
DC capa$_r^{2022Q3,0/1}$	0.66***		
	(15.8)		
DC capa$_r^{2022Q3}$		1.94***	
		(5.70)	
Log(1+DC capa$_r^{2022Q3}$)			0.77***
			(19.0)
Log(Plant capa$_{r,t}$)	0.051***	22.7***	0.18***
	(7.79)	(2.59)	(5.69)
Share RE/BESS$_{r,t}$	-0.23**	-163*	-0.23
	(-2.52)	(-1.78)	(-0.52)
Observations	741	741	741
Adj. R^2	0.32	0.66	0.49
Region FE	no	no	no
Year-Quarter FE	no	no	no

This table shows the relationship between the data center capacity in a region before and after the release of ChatGPT. The analysis is conducted at the region level. The sample period is from 2016 to 2025. In Column 1, the dependent variable is $DC\,capa_r^{\geq 2023Q1,0/1}$, which equals one if any new data center capacity is added after 2023Q1 in region r. The dependent variable in Column 2 is $DC\,capa_r^{\geq 2023Q1}$, which measures the new data center capacity that was added after 2023Q1 in region r, and Column 3 measures the logarithm of one plus this variable. $DC\,capa_r^{2022Q3}$ is the capacity of existing and planned data centers in region r as of September 30, 2022 (before the release of ChatGPT). t-statistics based on robust standard errors are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Appendix C: Robustness Tests

		All		RE/BESS		Gas	
Baseline specification		0.082***	(3.45)	0.075***	(3.19)	0.0052	(0.56)
Treatment	DC capa$_r^{2022Q3}$	0.000075***	(3.97)	0.000071***	(4.14)	0.000012	(1.22)
	Log(1+DC capa$_r^{2022Q3}$)	0.014***	(3.94)	0.013***	(3.79)	0.00066	(0.50)
	Top 20%	0.061***	(3.47)	0.064***	(3.57)	-4.9e-07	(-0.00)
	Top 5%	0.11***	(2.97)	0.10***	(2.79)	0.012	(0.84)
	$>$ 0 MW	0.061***	(3.47)	0.064***	(3.57)	-4.9e-07	(-0.00)
	$>$ 100 MW	0.079***	(3.47)	0.071***	(3.14)	0.0054	(0.60)
	Unadjusted treatment	0.066***	(2.60)	0.066**	(2.48)	-0.0046	(-0.51)
	Number of DCs	0.0027***	(3.94)	0.0026***	(4.47)	0.00045	(1.14)
Timing	Symmetric: $\pm$ 13 qtrs	0.070***	(3.06)	0.064***	(2.91)	-0.0024	(-0.23)
	Symmetric: $\pm$ 10 qtrs	0.071***	(2.91)	0.066***	(2.79)	-0.0066	(-0.59)
	Exclude 2022/23	0.11***	(3.91)	0.099***	(3.64)	0.014	(1.03)
	Year-level regressions	0.35***	(3.53)	0.31***	(3.18)	0.029	(0.74)
Box size	50 km	0.034***	(3.14)	0.033***	(3.12)	0.00035	(0.084)
	75 km	0.037**	(2.11)	0.037**	(2.12)	0.0053	(0.96)
	50 miles	0.070***	(4.37)	0.060***	(3.74)	0.0070	(1.07)
	100 miles	0.087***	(2.70)	0.071**	(2.11)	0.035*	(1.88)
Box size	200 km	0.082**	(2.40)	0.073**	(2.37)	0.0021	(0.28)
(continuous	300 km	0.10	(1.29)	0.088	(1.20)	0.010	(0.51)
investment)	200 miles	0.25***	(2.63)	0.20**	(2.51)	0.010	(0.47)
	300 miles	0.38*	(2.01)	0.30*	(1.76)	0.021	(0.41)

This table shows, for various alternative specifications, the coefficient estimates for $Post_t \times Treat_r$. The analysis is conducted at the region level. The sample period is from 2016 to 2025 (unless indicated otherwise). The dependent variable is $Capa_{r,t}^{0/1}$, which equals one if any new power plant is announced in year-quarter t and region r. The regression models follow the baseline specification in Table 2, Column 1. t-statistics based on robust standard errors clustered by regions are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

References

ACER, 2026, Agency for the Cooperation of Energy Regulators: Bidding zone review, https://www.acer.europa.eu/electricity/market-rules/capacity-allocation-and-congestion-management/bidding-zone-review, Accessed: 2026-03-13.

Bauer, Michael, Daniel Huber, Eric Offner, Marlene Renkel, and Ole Wilms, 2024, Corporate green pledges, Working Paper, SSRN 5052451.

BloombergNEF, 2026, Corporate clean energy buying fell in 2025 after nearly a decade of growth, https://about.bnef.com/insights/clean-energy/corporate-clean-energy-buying-fell-in-2025-after-nearly-a-decade-of-growth/, Accessed: 2026-03-15.

Bolton, Patrick, and Marcin Kacperczyk, 2026, Firm commitments, *Management Science* 72, 2134–2167.

Borenstein, Severin, James B. Bushnell, and Frank A. Wolak, 2002, Measuring market inefficiencies in California's restructured wholesale electricity market, *American Economic Review* 92, 1376–1405.

Brancucci, Cinzia, Juan Camus, Kevin Chu, Michael Coughlin, Dev Millstein, Nicholas Muller, Daniel O'Connell, and Michael Ting, 2025, Flexible data centers: A faster, more affordable path to power, Technical report, Princeton ZERO Lab.

Brannon, Ike, and Samuel Wolf, 2024, The impact of data centers on energy demand and market prices, Working Paper, SSRN 5017484.

Campello, Murillo, Chen Lin, Yue Ma, and Hong Zou, 2011, The real and financial implications of corporate hedging, *Journal of Finance* 66, 1615–1647.

Chen, Jiafeng, and Jonathan Roth, 2024, Logs with zeros? some problems and solutions, *Quarterly Journal of Economics* 139, 891–936.

Cornaggia, Jess, 2013, Does risk management matter? evidence from the U.S. agricultural industry, *Journal of Financial Economics* 109, 419–440.

Darmouni, Olivier, Clemens Lehner, and Pari Sastry, 2026, Financing investment in electricity, Working Paper.

Eldridge, Brent, Richard P. O'Neill, and Anya Castillo, 2017, FERC technical report on loss estimation: Marginal loss calculations for the dcopf, Technical report, Federal Energy Regulatory Commission, Washington, DC, January 24, 2017.

Ellison, Glenn, and Edward L. Glaeser, 1997, Geographic concentration in U.S. manufacturing industries: A dartboard approach, *Journal of Political Economy* 105, 889–927.

Esty, Benjamin C., and William L. Megginson, 2003, Creditor rights, enforcement, and debt ownership structure: Evidence from the global syndicated loan market, *Journal of Financial and Quantitative Analysis* 38, 37–59.

Feher, Adam, Emilia Garcia-Appendini, and Roxana Mihet, 2025, Is AI trained on public money? Evidence from US data centers, Working Paper, SSRN 5475309.

FERC, 2026, Form No. 714—Annual electric balancing authority area and planning area report: Overview, https://www.ferc.gov/industries-data/electric/general-information/electric-industry-forms/form-no-714-annual-electric/overview, Accessed: 2026-03-13.

Froot, Kenneth A., David S. Scharfstein, and Jeremy C. Stein, 1993, Risk management: Coordinating corporate investment and financing policies, *Journal of Finance* 48, 1629–1658.

Gilje, Erik P., and Jérôme P. Taillard, 2017, Does hedging affect firm value? Evidence from a natural experiment, *Review of Financial Studies* 30, 4083–4132.

Guidi, Gianluca, Francesca Dominici, Jonathan Gilmour, Kevin Butler, Eric Bell, Scott Delaney, and Falco J. Bargagli-Stoffi, 2024, Environmental burden of United States data centers in the artificial intelligence era, Working Paper.

Heiss, Alexander, and Thomas Schmid, 2026, Real effects of hedging on investment—The role of project finance, Working Paper.

Hogan, William W., 1992, Contract networks for electric power transmission, *Journal of Regulatory Economics* 4, 211–242.

Hui, Xiang, Oren Reshef, and Luofeng Zhou, 2023, The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market, Working Paper, SSRN 4527336.

International Energy Agency, 2025, Energy and AI, Technical report, International Energy Agency (IEA), Paris.

ISO New England, 2026, FAQs: Locational marginal pricing, https://www.iso-ne.com/participate/support/faq/lmp, Accessed: 2026-03-13.

Kay, Owen, Lutz Kilian, and Reid Taylor, 2026, Data center boom expected to raise electricity component of pce inflation, Working Paper, Federal Reserve Bank of Dallas.

Lin, Chen, Thomas Schmid, and Michael S Weisbach, 2021, Product price risk and liquidity management: Evidence from the electricity industry, *Management Science* 67, 2519–2540.

Mamkhezri, Jamal, Xiaochen Sun, and Yuting Yang, 2025, The hidden cost of the cloud: Data centers and electricity market inefficiency, Working Paper, SSRN 5736562.

Masanet, Eric, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan G. Koomey, 2020, Recalibrating global data center energy-use estimates, *Science* 367, 984–986.

Muir, Martha, 2025, Are data centres a setback for the green energy transition?, https://www.ft.com/content/6e4193a0-aaf3-4752-8b45-696eae57ba18, Accessed: 14 March 2026.

Mytton, David, Hannah Ritchie, Benjamin K. Sovacool, and Mariana Mazzucato, 2023, The stretched grid: Can increasing electrification and demand for data centers be met without overhauling the electricity system?, *Oxford Open Energy* 2.

Pérez-González, Francisco, and Hayong Yun, 2013, Risk management and firm value: Evidence from weather derivatives, *Journal of Finance* 68, 2143–2176.

PJM Interconnection, 2023, Dominion data center: Table B9 in 2023 load forecast, https://www.pjm.com/-/media/DotCom/committees-groups/subcommittees/las/2023/20230626/20230626-item-03---table-b9-in-load-report.ashx, Accessed: 2026-03-11.

Ryan, Nicholas, 2021, Holding up green energy, Working Paper, SSRN 3905718.

S&P Global, 2026, 451 research – products and solutions, https://www.spglobal.com/energy/en/products-solutions/451-research, Accessed: 13 March 2026.

Texas Comptroller of Public Accounts, 2024, Data centers, https://comptroller.texas.gov/taxes/data-centers/, Accessed: 2026-03-15.

U.S. Department of Energy, Grid Deployment Office, 2023, National transmission needs study, https://www.energy.gov/gdo/national-transmission-needs-study, Accessed: 2026-03-11.

Virginia JLARC, 2025, Data centers in Virginia, Report, Virginia General Assembly, Virginia Joint Legislative Audit and Review Commission.

Volcovici, Valerie, and Laila Kearney, 2024, How ai and cloud computing may delay the transition to clean energy, https://www.reuters.com/technology/artificial-intelligence/how-ai-cloud-computing-may-delay-transition-clean-energy-2024-11-21/, Accessed: 2026-03-15.