Who Pays for Growth? Evidence from Data Centers and the Grid*

Robyn Meeks† Jacquelyn Pless*‡ Zhiyuan Qi§ Zhenxuan Wang¶

October 23, 2026

Abstract

Who bears the costs when technological change requires new infrastructure? By linking U.S. data center entry from 2010 to 2024 to electric utilities' prices, investments, and regulatory outcomes, we find that data center entry raises average retail rates by 2.7% and residential retail rates by 2.1%. Effects are concentrated among investor-owned utilities (5.6% on average) and are larger in restructured electricity markets. Infrastructure investment and utilities' authorized rate bases increase, but price effects are lower where utilization improves. These findings suggest that the incidence of infrastructure-intensive technological change depends on system conditions and the institutions governing investment and cost recovery.

Keywords: data centers; incidence; infrastructure investment; electricity prices; technological change; rate-of-return regulation
JEL Codes: Q41, L94, H23, L51, O33

*For helpful comments and suggestions, the authors thank Steve Cicala, Joshua Gans, Paul Joskow, Hong Luo, and participants of presentations at Berkeley Energy Camp, the Toronto Meeting on the Economics of Climate Change, LSE Environment Week, UT-Austin, Trinity University, East Carolina University, SWEEEP, and the IADB. Zhenxuan Wang gratefully acknowledges funding support from the Institute for Humane Studies (grant no. IHS020420) for data purchase.

†Sanford School of Public Policy, Duke University. Email: robyn.meeks@duke.edu.

‡*Corresponding author. Rotman School of Management, University of Toronto. Email: jacquelyn.pless@rotman.utoronto.ca.

§Department of Economics, North Carolina State University. Email: zqi8@ncsu.edu.

¶Department of Agricultural and Resource Economics, North Carolina State University. Email: zhenxuan.wang@ncsu.edu.

1 Introduction

Artificial intelligence (AI) is unleashing one of the largest infrastructure booms in modern history. Behind the rapid advances in frontier AI models, cloud computing, and digital services lies an enormous and fast-growing physical footprint: hyperscale data centers consuming vast quantities of electricity. Global investment in data centers that rely crucially on a stable source of electricity supply nearly doubled between 2022 and 2024 (IEA, 2025), reaching roughly half a trillion dollars, and projections indicate that electricity use from data centers in the U.S. alone could more than double by 2028 (Smith et al., 2026). The scale and speed of this expansion have transformed data centers into one of the fastest-growing sectors of the U.S. economy and the most consequential source of electricity demand growth in the twenty-first century.

This surge arrives after decades of relatively flat U.S. electricity demand and represents a fundamentally different form of load growth from what electricity systems have historically been built to accommodate. Unlike traditional demand growth—which tends to be gradual, diffuse, and tied to population or economic expansion—data centers arrive in large, concentrated increments and operate at high utilization rates. Their geographic concentration can also increase the infrastructure demands placed on individual utilities and local electricity systems (Fang and Greenstein, 2025). Serving these loads may require substantial investment in electricity generation, transmission, and distribution infrastructure, especially in areas where the grid is already constrained, and concerns over how this might impact ratepayers' electricity bills have become central to debates over data center expansion.[1] But the consequences for retail rates are not mechanically determined by the size of the load shock, nor are they determined by a market-clearing price. They depend on whether new load improves system utilization, how utility ownership and electricity industry structure shape investment incentives and cost recovery, and whether costs are assigned directly to data centers or socialized across other customers.

In this paper, we ask: who pays for (load) growth? We study how data center entry affects local retail electricity rates in the U.S., with particular attention to how these effects differ across residential, commercial, and industrial customers and across the institutions that govern rate-making. When systems have sufficient spare capacity and customers with high load factors (like data centers) contribute to fixed cost recovery, data centers may have little effect on average rates or even could put downward pressure on rates. When systems are capacity-constrained, require upgrades or expansions, or rely on costly wholesale procurement, though, large new loads may place upward pressure on retail rates. This ambiguity is a general feature of load growth—additional sales can spread fixed costs over more kilowatt-hours, but it can also increase the need for costlier generation and grid infrastructure investment (Wiser et al., 2025). A utility's

[1] Concerns around rapid expansion of data centers—especially the implications for electricity bills—have received extensive national attention (e.g., see WashingtonPost (2026), Bloomberg (2026), and Reuters (2026)). Data center opposition has been referred to as "the most bipartisan issue since beer" (NYTimes, 2026), and at least 15 states have proposed or adopted data center moratoriums or bans (NCSL, 2026).

ownership model also impacts its investment incentives and cost recovery mechanisms. Moreover, electricity industry structure can impact generation costs and markups, but those differences do not translate one-for-one into retail prices given how utilities recover costs through rate base, tariffs, and delivery charges (Cicala, 2022, 2025; MacKay and Mercadal, 2026).

A key contribution of this paper is to estimate the retail incidence of infrastructure-intensive demand growth and to trace these effects to the underlying mechanisms through which utilities accommodate growth and recover costs. To do so, we link facility-level data center entry across the U.S. from 2010 through 2024 to utility-state-level retail electricity prices, combining data on the location, timing, and scale of data centers with utility-level retail rates by customer class, utility ownership, and electricity industry structure. We also assemble utility-level data on infrastructure investments, operational outcomes, regulatory rate cases, large-load tariffs, and pre-existing system conditions.

This focus distinguishes our analysis from contemporaneous related work on location choice and dispatch model estimates of wholesale price and emissions impacts (Elhai and Zhou, 2026), state-level average retail rate effects (Watten, Bistline and Blanford, 2026), household price effects during the post-AI build-out (Scalera, Bosetti and Pecci, 2026), and the broader local economic effects of data centers (Alvarez et al., 2026; Yue and Zeng, 2026). Instead, we study retail rate incidence at the level where rates are set (utility), distinguish between households and firms, and focus on the heterogeneity across institutions that influences utilities' investment decisions, cost recovery, and rate-making.

Our empirical approach uses the Callaway and Sant'Anna (2021) estimator and compares changes in outcomes for utility-state pairs whose service territories receive a data center to those that never do. This design allows us to assess pre-trends and examine how retail rate effects evolve over time. One challenge is that data centers do not locate randomly. Firms select locations based on factors that may also affect electricity prices, such as electricity access, digital connectivity, grid capacity, land availability, permitting, and tax incentives. We address this concern in several ways. Our baseline design absorbs time-invariant differences across utility-state pairs and common annual shocks. We also show robustness to allowing utilities and regions with different baseline characteristics to follow different time paths, to dropping potentially contaminated control units within multi-state utilities, and to a matched difference-in-differences design that balances never-treated utilities to treated utilities on pre-period electricity prices, customer counts, manufacturing intensity, and income. As a complementary analysis, we also estimate instrumental variable specifications that use interactions between historical local exposure to the data processing sector and national-level capital expenditures by large cloud computing firms to estimate the effects of cumulative data center capacity.

We find that, on average, retail electricity rates increase following data center entry, but the effects differ across customer segments, utility ownership type, and electricity industry structure. On average, retail prices rise by 2.7%. Residential rates increase by 2.1%, commercial rates by

2.8%, and industrial rates by 4.2%. Given our research design, these increases are attributable to all data center growth following an initial entry rather than to a single facility alone. The positive residential effect indicates that at least some of the costs associated with data center growth are passed through to households. The effects for commercial and industrial customers suggest that non-residential customers also face price increases, but since data centers themselves may be recorded as commercial, industrial, special contract, or large-load customers, these estimates embed both pass-through to incumbent firms and the prices paid by data centers. The estimates overall remain similar across numerous robustness checks, are similar or larger in the matched design, and are supported by capacity-based IV estimates.

Given how utility ownership determines the incentives and governance structures through which infrastructure costs are financed and recovered, we next examine heterogeneity by ownership model. Investor-owned utilities (IOUs), which serve more than 70% of U.S. electricity customers, are shareholder-owned firms that recover approved costs through regulated rates and earn allowed returns on capital included in rate base. Public utilities and cooperatives, by contrast, are owned by governments or members. Therefore, they may place different weight on rate stability, local economic development, and avoiding cross-subsidies. We find that retail price increases are concentrated among IOUs, for which data center entry increases average rates by 5.6%, with residential prices rising by 5.2%, commercial prices by 4.6%, and industrial prices by 8%. Effects are much smaller among public utilities and absent among cooperatives.

We also examine heterogeneity by electricity industry structure, because the same load shock may be absorbed differently depending on whether the utility owns and plans generation (as in non-restructured states characterized by vertically integrated utilities) or instead procures power through market-based arrangements (as in restructured states). In states with restructured electricity industries, generation has been separated from transmission and distribution, and electricity supply is procured.[2] These institutional differences can affect both the costs created by data center load and the channels through which those costs enter retail rates. We find larger effects for utilities operating in restructured states, with average retail prices rising by 3.8%, compared to a smaller and statistically insignificant increase in non-restructured states. However, a two-way split suggests that ownership is the first-order source of heterogeneity, as price effects are concentrated among IOUs in both restructured and non-restructured states.

Translating these estimates into impacts on household electricity bills, the estimated 2.1% increase in residential prices corresponds to an average increase of about $38.52 per household per year, or roughly $3.21 per month. This is relative to a pre-treatment annual residential electricity bill of about $1,835 per year, or $153 per month. For residential customers served by IOUs, the 5.2% increase in retail rates translates into an additional $7 per month relative to a pre-treatment average monthly bill of $141, or $88 per year.

We next provide evidence on mechanisms, focusing on IOUs, the set of utilities for which price

[2]Procurement comes through wholesale markets, default service agreements, contracts, or retail suppliers.

effects are largest and for which we have investment and rate case data. We find that data center entry induces real infrastructure responses. Electricity-specific construction activity, generation investment, distribution equipment (in restructured states), and authorized rate bases increase. These results indicate that utilities invest and capital investment enters the regulatory cost-recovery process following data center entry.

But investment does not translate uniformly into higher retail rates given how price effects differ for restructured and non-restructured states. We investigate this further by examining operational responses. In non-restructured states, load factors improve, consistent with utilities absorbing part of the new load through higher utilization, owned generation, and gradual cost recovery. On the other hand, in restructured states, utilities exhibit earlier construction-stage capital activity and no comparable utilization offset. They also show more active cost-recovery responses, including more rate case and rider proceedings, higher requested revenue requirements, and larger customer advances for construction. These patterns are consistent with retail pass-through being greatest when new load creates system cost pressure that is not offset by utilization gains.

Additional results showing how the effects vary based on system constraint proxies reinforce this interpretation. Price effects are larger in low-reserve-margin regions, particularly in restructured states, consistent with a capacity-adequacy mechanism. That is, when systems have limited spare capacity, large new loads are less likely to be absorbed through existing or planned capacity and more likely to require costly procurement, reliability expenditures, or accelerated infrastructure investment. We also find larger effects in fossil-heavy restructured regions. Wholesale market exposure also appears to contribute to cost pressure but does not fully explain the retail rate effects. Data center entry raises the congestion and loss components of wholesale prices rather than the energy component, consistent with localized network stress. Controlling for observed wholesale prices attenuates the retail price effects slightly but does not eliminate them.

Taken together, these findings shed light on how the retail incidence of data center growth is mediated through institutions and system conditions. A utility's business model appears to be central, with price effects primarily concentrated among IOUs rather than public utilities or cooperatives. Among IOUs, data centers can raise rates when incremental system costs exceed the utilization benefits of new load and when those costs are socialized through retail rates. But effects can be muted where utilization improves. This ambiguity is consistent with broader evidence that load growth can affect retail prices through offsetting channels, including fixed-cost spreading, costlier generation, and grid infrastructure investment (Wiser et al., 2025). Moving forward, cost-causation mechanisms like large-load tariffs may also mute price effects by assigning incremental costs directly to the facilities that create them, but that is left for future research.

For policymakers and regulators evaluating proposed data centers, the central question is not simply how large a facility is or how much electricity it consumes. The relevant questions are whether infrastructure costs are assigned to the customer causing them and whether the utility and market institutions are likely to socialize remaining costs across incumbent ratepayers.

Future effects could be larger if new facilities locate in more constrained systems with longer interconnection queues, equipment shortages, or costly procurement needs. On the other hand, they could be smaller, zero, or negative if very large high-load-factor customers generate substantial revenue, spread fixed costs over a larger sales base, or are increasingly covered by specialized large-load tariffs and direct cost-assignment mechanisms.

Related Literature and Contributions. This paper contributes to a rapidly growing literature on the economic consequences of data centers. Several contemporaneous papers study related questions but focus on different outcomes, units of analysis, and mechanisms. Alvarez et al. (2026) and Yue and Zeng (2026) examine the effects of data center entry on local economic activity like employment, wages, housing markets, and average electricity prices. Kay, Reaser and Taylor (2026) and Mamkhezri, Sun and Yang (2025) find that data center growth can increase *wholesale* electricity prices (as opposed to the retail rates that end-users face) and generation costs in regions facing constraints. Elhai and Zhou (2026) combine dispatch and data center location choice models to study wholesale price and emissions effects, showing that data centers tend to locate in regions with flatter electricity supply curves, so wholesale price effects need not lead to retail price changes. Scalera, Bosetti and Pecci (2026) study household electricity prices in utility territories during the post-2021 AI-associated buildout, finding that data centers increase residential prices, with smaller effects in non-profit utility territories and places with energy surplus. Watten, Bistline and Blanford (2026) estimate retail rate effects at the state level and find modest average price decreases, interpreting the result through economies of scale and fixed cost spreading.[3] Others focus on the environmental consequences of data centers (Knittel, Senga and Wang, 2025; Muller, 2026; Hernandez-Cortes, Meng and Weber, 2026).

Our analysis differs from these papers in both the object of interest and the mechanisms. We focus on local retail rate incidence: whether the costs of serving new data center load are socialized through the average rates paid by incumbent households and firms. We work at the utility-state level—the level where retail rates are set. We estimate effects separately by customer class, distinguish important heterogeneity across utility ownership models, and compare effects for utilities in restructured and non-restructured markets. We then connect the price effects to direct evidence on mechanisms. Our contribution thus is not only to estimate whether data centers have affected retail prices, but to explain when those effects have been positive, muted, or offset by utilization gains, and through which institutions the costs are allocated.

Our paper also contributes to broader literatures in regulatory, energy, and public economics. Classic work on cost-of-service emphasizes how regulation shapes utility investment behavior, cost-minimization incentives, and capital expansion (Averch and Johnson, 1962; Joskow, 1974;

[3] Watten, Bistline and Blanford (2026) find that data centers reduce state-level average retail rates. However, their preferred estimates rely on comparisons across states without state fixed effects and a shift-share instrumental variable constructed as the interaction between an indicator for post-2021 years, national data center construction spending, and a utility territory's proportion of national pre-2021 fiber infrastructure. Moreover, state averages ignore the local and concentrated nature of data center entry. Costs are recovered and allocated at the utility level.

Joskow and Rose, 1989; Joskow and Schmalensee, 2025). A large literature studies market power, procurement, and pricing in restructured electricity markets (e.g., Wolfram, 1999; Davis and Wolfram, 2012; Fabra and Reguant, 2014; Cicala, 2015). Recent studies of modern electricity markets revisit these issues, showing that regulated returns, rate case incentives, and restructuring can shape capital investment, asset retirement, and consumer costs (Cicala, 2025; Dunkle Werner and Jarvis, 2026; Gowrisankaran, Langer and Reguant, 2026). We study how a new, consequential source of load growth impacts retail rates and how those effects differ by utility ownership, electricity industry structure, and regulatory institutions.

Our paper also speaks to the literature specifically on cost pass-through in energy markets. Recent work shows that restructuring can reduce generation costs, but that cost savings may not be passed through fully to consumers once retail prices, markups, and regulated delivery charges are considered (Cicala, 2022; MacKay and Mercadal, 2026; Cicala, 2025). Others examine how energy and environmental cost shocks—like emissions costs and energy input costs—are shifted through markets (Fabra and Reguant, 2014; Ganapati, Shapiro and Walker, 2020).

Our study also complements work showing that physical grid constraints and transmission institutions shape electricity market outcomes. Transmission constraints can limit gains from trade, raise generation costs, and create distributional conflicts (Davis and Hausman, 2016; Hausman, 2025). We show that retail price effects are larger in low-reserve-margin regions and fossil-heavy restructured markets, consistent with how system constraints matter for whether a region can absorb new load and the retail incidence of demand growth. Less directly, our paper contributes to the literature on the costs and benefits of infrastructure investments (e.g., Donaldson (2018) and Allen and Arkolakis (2022)), which emphasizes how infrastructure expands market access, raises productivity, and reshapes economic activity. Our analyses focus on who pays for new infrastructure required to meet demand induced by technological change.

Finally, the paper contributes to the literature on public versus private ownership. Classic theory emphasizes that ownership affects cost-reduction incentives, investment, and quality (Hart, Shleifer and Vishny, 1997; Shleifer, 1998). Electricity has long been a central setting for studying public versus private ownership, with work comparing prices and efficiency across utility ownership models (Färe, Grosskopf and Logan, 1985; Atkinson and Halvorsen, 1986; Kwoka, 2005).[4] We extend this debate to the incidence of infrastructure-intensive demand growth.

Outline. Section 2 provides background on data center growth, electricity markets, and utility regulation. Section 3 provides a conceptual framework. Sections 4 and 5 describe our data and empirical strategies. Section 6 presents the main price results and institutional heterogeneity. Section 7 examines mechanisms. We conclude in Section 8.

[4]Empirical work has also studied these tradeoffs in public service settings such as health care and health insurance (Duggan, 2000; Duggan, Gruber and Vabson, 2018; Duggan et al., 2023).

2 Background

This section provides institutional background on data center load growth, utility ownership, retail rate setting, and electricity market restructuring that is most relevant for our paper. Additional details can be found in Appendix A.

2.1 Data Centers, Electricity Demand, and Grid Investment

Data centers are energy-intensive facilities that operate continuously to support cloud computing, digital services, and artificial intelligence. Their electricity demand is driven by both the computing equipment itself as well as cooling requirements. Electricity use from data centers in the U.S. roughly doubled between 2000 and 2005, then remained comparatively flat through the early-to-mid 2010s despite continued growth in digital activity, largely because of efficiency improvements in servers, cooling, computation, power management, and utilization (Masanet et al., 2020; Shehabi et al., 2024). That period of flat electricity use has ended. U.S. data center electricity use more than doubled between 2017 and 2023, reaching about 176 TWh, or 4.4% of total U.S. electricity consumption, and projections indicate that data centers could account for 6.7% to 12% of U.S. electricity consumption by 2028 (Shehabi et al., 2024).

This demand growth creates a unique challenge for the electricity sector. Unlike traditional load growth, which is often gradual and geographically diffuse, data centers tend to arrive in large, concentrated increments and operate at high utilization rates. Serving this load may require utilities or wholesale markets to procure additional generation, expand transmission and distribution infrastructure, upgrade substations and transformers, and undertake reliability-related investments. These pressures can be especially large when data centers locate in regions with limited spare capacity or when multiple facilities cluster in a small number of service territories (Satchwell et al., 2025; Fang and Greenstein, 2025).

Planning for data center load is also difficult because the timing and scale of realized demand can depart from forecasts. Utilities may need to begin engineering, interconnection, and infrastructure investments long before the full load materializes. Moreover, if expected demand does not arrive, or if a customer later exits or scales down, utilities could be left with long-lived assets whose costs cannot be fully recovered from the customer that induced the investment requirement. These risks have made the allocation of data center-related infrastructure costs a central concern for utilities, regulators, and ratepayers.

2.2 Utility Ownership and Governance

Electric utilities in the U.S. differ in ownership and governance in ways that shape incentives, financing, and cost recovery. We group utilities into three categories: investor-owned utilities, public utilities, and cooperatives. Investor-owned utilities (IOUs) are for-profit firms owned

by shareholders. They are typically subject to state public utility commission regulation for retail distribution service and, where relevant, federal regulation for interstate transmission and wholesale power sales. Public utilities, including municipal utilities and other government-owned utilities, are owned by public entities and typically governed by city councils, public utility boards, or similar local authorities. Cooperatives are member-owned utilities governed by customer-elected boards and generally operate on a not-for-profit basis.

IOUs recover approved costs through regulated rates and earn allowed returns on capital investments included in rate base. Public utilities and cooperatives do not distribute profits to shareholders and finance investments through municipal debt, cooperative financing arrangements, member equity, or other mechanisms. Their governing bodies may also place different weights on rate stability and local economic development. As a result, the same physical load shock may translate into different investment responses and different retail price effects depending on utility ownership.

2.3 Retail Rate Setting and Cost Recovery

Retail electricity rates are not determined by a market-clearing price. For regulated utilities, rates are set through administrative processes designed to allow utilities to recover prudently incurred costs, and IOUs earn an allowed return on approved capital investments. The central component determined in this process is the revenue requirement: the amount of revenue the utility is authorized to collect from customers.

In a general rate case, the regulator first determines the utility's revenue requirement. This calculation includes operating expenses, depreciation, taxes and fees, and the allowed return on the utility's rate base. The rate base is the stock of approved long-lived assets used to provide service, such as generation assets, poles, wires, substations, transformers, and other utility infrastructure. To be included in rate base, investments generally must be found prudent and used and useful. Operating expenses, including labor, fuel, and purchased power, are typically recovered on a dollar-for-dollar basis.

After determining the revenue requirement, regulators allocate costs across customer classes, typically including residential, commercial, industrial, and sometimes other tariff classes. This allocation is based on cost-of-service studies and cost-causation principles, but it also involves administrative judgment. Costs may be allocated according to customer counts, energy consumption, contribution to peak demand, voltage level, load factor, and other service characteristics. The final step is rate design: translating each class's allocated revenue requirement into prices and charges, such as volumetric energy charges, demand charges, customer charges, or special tariff provisions. Public utilities and cooperatives use different governance processes, but they face the same basic problem of allocating common system costs across customer groups.

Cost recovery need not occur only through full general rate cases. Because rate cases occur

periodically, utilities may face regulatory lag between when costs are incurred and when rates are adjusted. Several mechanisms can reduce this lag. Fuel adjustment clauses and purchased-power adjustments allow some energy costs to pass through between rate cases. Infrastructure riders or trackers allow recovery of specific categories of capital spending. Construction work in progress (CWIP) can allow utilities to earn on unfinished projects before they enter service, while allowance for funds used during construction (AFUDC) capitalizes financing costs during construction and adds them to rate base once the asset enters service. These mechanisms matter for data centers because large-load growth may require infrastructure investment before new load is fully realized.

2.4 Electricity Industry Structure and Wholesale Exposure

Electricity market structure also can impact rate-making. Some states have what we refer to as "restructured" electricity industries, whereby generation is unbundled from transmission and distribution. This is distinct from whether a state allows for retail choice such that customers choose their electricity supplier. In traditionally regulated states (henceforth "non-restructured markets"), vertically integrated utilities often own or procure generation, transmission, and distribution as part of an integrated planning process. Their rate base may include generation, transmission, and distribution assets, and their operating expenses may include fuel and purchased power.

In "restructured" states, the regulated utility is often primarily a delivery utility. Generation is supplied through wholesale markets, default-service procurement, contracts, competitive suppliers, or other market-based arrangements, while the delivery utility continues to recover regulated transmission and distribution costs through retail rates. These utilities therefore may be more exposed to changes in wholesale energy prices, capacity prices, congestion, ancillary services, and default-service procurement costs. At the same time, restructuring should not be conflated with participation in an RTO or ISO: some vertically integrated utilities participate in organized wholesale markets, and transmission cost recovery may be governed by FERC-regulated mechanisms even in otherwise traditionally regulated states.

This distinction matters for data center growth. In non-restructured settings, utilities may be able to smooth load growth through owned generation, integrated resource planning, long-term procurement, reserve margins, and gradual rate recovery. In restructured settings, large new loads may translate more directly into procurement, capacity, congestion, or default-service costs, while the delivery utility continues to recover transmission and distribution investments through regulated rates.

3 Conceptual Framework

We develop a simple framework to clarify how data center entry can increase, decrease, or leave unchanged retail electricity prices, and why these effects may differ across customer class, utility ownership type, and electricity industry structure. We begin with the standard cost-of-service

model of utility regulation and then introduce data center load, direct cost assignment, utilization, customer class allocation, and institutional heterogeneity.

3.1 Cost-of-Service Regulation and Average Prices

Following Joskow and Schmalensee (2025), under cost-of-service regulation (COSR), a utility's revenue requirement (or cost of service) in year t can be written as

$$RR_t = O_t + D_t + r_t RB_t + F_t, \tag{1}$$

where O_t denotes operating and maintenance expenses, D_t is annual depreciation, RB_t is the regulated rate base, r_t is the allowed rate of return on the rate base, and F_t captures other costs such as taxes and fees. The product $r_t RB_t$ is the allowed return on rate base, or the capital charge component of the annual revenue requirement. For our setting, it is also useful to distinguish between energy-related and non-energy operating expenses:

$$O_t = O_t^N + M_t, \tag{2}$$

where O_t^N captures non-energy operating expenses, such as engineering, maintenance, interconnection studies, and reliability planning, while M_t captures energy-related costs like fuel, purchased power, wholesale energy, capacity, congestion, and default-service procurement costs. This decomposition simply separates the component of operating costs related to generation and wholesale market conditions from other operating costs.

Let baseline electricity sales prior to data center entry be Q_0, so that the average retail price is

$$p_0 = \frac{RR_0}{Q_0}. \tag{3}$$

This price measure summarizes utilities' revenue required per kilowatt-hour sold, providing a basis for considering how data center load shocks might impact average retail rates. In practice, customers face class-specific tariffs and alternative rate designs (e.g., time-of-use rates), and in our empirical analyses, we construct average prices as total *observed* revenue over total electricity sales.

3.2 Benchmark Case: Data Center Entry and Average Price Effects

Suppose a data center enters a utility's service territory with expected annual electricity consumption x. Let the change in the utility's annual recoverable revenue requirement induced by serving the data center be written and decomposed as:[5]

[5]We use Δ notation to capture changes because data center entry is a discrete event that may affect planning, infrastructure investment, regulatory proceedings, and cost recovery before realized electricity consumption is fully observed.

$$\Delta RR(x) = \Delta O^N(x) + \Delta M(x) + \Delta D(x) + \Delta(rRB)(x) + \Delta F(x). \tag{4}$$

These components correspond to the main mechanisms we examine empirically. That is, data center entry may increase energy and non-energy operating expenses as well as capital investment in generation, transmission, and distribution assets, which affects the revenue requirement through the rate base and depreciation.

As a benchmark, suppose all incremental recoverable costs are socialized across a utility's entire customer base. The post-entry average price is then

$$p_1^B = \frac{RR_0 + \Delta RR(x)}{Q_0 + x}, \tag{5}$$

where B denotes the benchmark case. Average retail prices increase if and only if

$$p_1^B > p_0 \quad \Longleftrightarrow \quad \frac{\Delta RR(x)}{x} > p_0. \tag{6}$$

Thus, data center entry raises average retail prices when the incremental recoverable cost of serving the new load exceeds the utility's pre-existing average cost of service. Prices may remain unchanged or fall if the data center can be served at low incremental cost relative to pre-entry average costs. For example, this could occur if the system has excess capacity to absorb the new load and utilization of existing assets improves, or the benefit of spreading fixed infrastructure costs over additional kilowatt-hours sold is large.

3.3 Direct Cost Assignment

Some costs may be assigned directly to the data center, such as through customer advances for construction, and more recently, through special contracts or large-load tariffs that specify contributions in aid of construction or minimum demand charges.[6] Let $\psi \in [0, 1]$ denote the share of data center-related incremental costs assigned directly to the data center. Let $\omega \in [0, 1]$ denote the share of the remaining incremental revenue requirement socialized across incumbent customers through general retail rates. Ordinary tariff payments from the new load can help recover existing common costs, spreading these costs over a larger sales base.

Let $U(x)$ denote the cost-spreading benefit associated with the data center, defined as their tariff contribution toward the pre-entry revenue requirement RR_0. The post-entry revenue requirement

[6]Costs also might not be entirely borne by the utility's retail customers if they are deferred because of regulatory lag, financed through municipal or cooperative governance mechanisms, or allocated regionally through transmission cost-sharing rules.

allocated to incumbent ratepayers is

$$RR_{\text{inc},1}(x) = \underbrace{RR_0 - U(x)}_{\text{Remaining pre-existing costs}} + \underbrace{\omega(1-\psi)\Delta RR(x)}_{\text{Allocated incremental costs}}. \tag{7}$$

Relative to the baseline, the net revenue pressure placed on incumbent ratepayers is

$$N(x) = RR_{\text{inc},1}(x) - RR_0 = \omega(1-\psi)\Delta RR(x) - U(x). \tag{8}$$

This expression reflects the incidence relationship of interest. Incumbent customers face upward rate pressure when the portion of incremental costs recovered through general rates exceeds the utilization and cost-spreading benefit of the new load. If $N(x) > 0$, data center entry creates upward pressure on retail rates. If $N(x) = 0$, the cost and utilization effects offset. If $N(x) < 0$, data center entry can lower rates.

3.4 Customer Class Cost Allocation

After the revenue requirement is established, regulators or governing boards allocate costs across customer classes and design tariffs to recover each class's assigned revenue requirement. Let $k \in \{R, C, I\}$ denote residential, commercial, and industrial customers. Let θ_k denote the share of net incumbent revenue pressure allocated to class k, where

$$\sum_{k \in \{R,C,I\}} \theta_k = 1. \tag{9}$$

Let Q_k denote baseline sales to incumbent customers in class k. The change in the average price paid by incumbent customers in class k is:

$$\Delta p_k^{inc} = \frac{\theta_k N(x)}{Q_k} = \frac{\theta_k \left[\omega(1-\psi)\Delta RR(x) - U(x)\right]}{Q_k}. \tag{10}$$

Class-specific effects thus depend on the net revenue pressure and the share allocated to each customer class. The allocation shares θ_k are determined by functionalized cost allocation rather than by market forces, with costs typically assigned to system functions—generation, transmission, distribution, and customer service—and then allocated across classes according to cost drivers, such as contribution to peak demand, energy use, customer counts, voltage level, load factor, and service characteristics. This matters for data centers because their voltage of interconnection and load shape may determine which system costs they are assigned.[7]

Historically, many data centers were classified under commercial tariffs, though some larger facilities may be classified as industrial, special-contract, or large-load customers depending on

[7] A large facility connected at sub-transmission or transmission voltage may bypass some distribution costs that residential and small commercial customers pay, while smaller facilities connected at distribution voltage may contribute to costs shared with residential and commercial customers.

size, voltage, and utility-specific tariff definitions. Observed commercial and industrial average prices may thus embed changes in prices paid by incumbent firms as well as the tariff paid by the data center itself. Residential price effects provide the cleanest evidence of pass-through to incumbent customers because data centers do not enter the residential class.

3.5 Heterogeneity by Utility Ownership

The parameters in the framework may differ systematically by utility ownership. For IOUs, approved investments enter rate base and earn an allowed return, creating an incentive to make capital investments. Let ρ denote the utility's effective cost of capital. The shareholder value associated with a data center-induced increase in rate base can be approximated as

$$\Delta V^{IOU}(x) \approx (r - \rho)\Delta RB(x) - \kappa(x), \tag{11}$$

where $\kappa(x)$ captures regulatory risk, disallowance risk, financing costs not reflected in the allowed return, and other costs. When $r > \rho$ and the investment is likely to be approved as prudent, used, and useful, rate base expansion can increase shareholder value. As such, IOUs have a strong incentive to accommodate data center load through capital projects.

On the other hand, public utilities and cooperatives do not distribute profits to shareholders, so the rate base incentive is absent or substantially muted. Their capital costs may instead reflect debt service, coverage ratios, municipal finance, member equity, or cooperative financing arrangements. Their governing boards may also place more direct weight on other factors, such as rate stability, local economic development, or avoiding cross-subsidies among customer types.

These institutional differences can appear in several ways. Public utilities and cooperatives may have lower $\Delta RR(x)$ if they finance capital at lower cost or pursue less capital-intensive responses. They may have lower ω if costs are financed outside general volumetric rates. They may have higher ψ if they require large customers to pay directly for dedicated infrastructure. Or they may have different θ_k if their governing boards allocate costs differently across customer classes. The empirical prediction is therefore not simply that public utilities or cooperatives face no costs from data center entry. Rather, it is that similar physical load shocks may translate into smaller or different retail price effects, since ownership changes the incentives and mechanisms for investment, financing, and cost recovery.

3.6 Heterogeneity by Electricity Industry Structure

Electricity industry structure provides a second important source of heterogeneity. A key difference is whether generation is owned and planned by the regulated utility (in non-restructured states) or procured through wholesale markets, default-service procurement, contracts, or retail suppliers (in restructured states).

For a traditionally regulated vertically integrated utility, the data center-induced change in the revenue requirement may include generation, transmission, and distribution assets owned by the utility, so they may face higher long-run capital needs. At the same time, they may be able to accommodate new load through owned generation, existing reserve margins, long-term procurement, and integrated resource planning.

In restructured states, regulated delivery utilities are less likely to own generation, so their rate bases may consist primarily of delivery assets, while energy supply is procured through wholesale markets, contracts, default-service auctions, competitive suppliers, or other mechanisms. The $\Delta(rRB)(x)$ component of the revenue requirement response may thus primarily reflect delivery assets, and $\Delta M(x)$ captures wholesale energy, capacity, congestion, ancillary services, and default-service procurement costs. That said, restructuring did not entirely eliminate utility-owned generation everywhere, so some utilities in restructured states retained generation assets.

The distinction between restructured and non-restructured states therefore changes both the composition of costs and the likely offsetting forces. In non-restructured states, data center entry may induce generation and network investment, but utilities may also smooth costs through planning, owned assets, long-term procurement, and utilization gains. Similar investments might emerge for utilities in restructured states, although data center entry may also generate more immediate cost pressure through wholesale energy, capacity, congestion, default-service procurement, and regulated delivery investment.[8]

3.7 Empirical Implications

The framework yields five implications that guide our empirical analyses. First, the sign of the retail price effect is theoretically ambiguous. Data center entry raises average prices when the incremental recoverable cost of serving the new load exceeds the pre-existing average cost of service, but it can leave rates unchanged or reduce them when utilization gains, ordinary tariff revenues, or direct cost assignment offset incremental costs.

Second, pass-through to incumbent customers depends on the net revenue pressure. Retail prices rise for incumbent customers when general rate-recovered costs exceed utilization and cost-spreading benefits. This distinction explains why infrastructure investment need not translate mechanically into retail price increases. Third, customer class incidence depends on administrative cost allocation and rate design. Residential, commercial, and industrial rates may respond differently because costs are functionalized and allocated based on cost drivers. Residential estimates provide the cleanest evidence of pass-through to incumbent households, while commercial and industrial averages may also reflect the tariffs paid by data centers themselves.

[8] Restructuring should not be conflated with RTO/ISO membership. Some vertically integrated utilities participate in organized wholesale markets, and some transmission costs in RTO/ISO regions are recovered through FERC-regulated formula rates. Thus, restructuring affects the generation and procurement margin, while RTO/ISO participation may affect the transmission cost-recovery margin.

Fourth, utility ownership should matter. IOUs may have stronger incentives to pursue capital investments because approved assets enter rate base and earn an allowed return. Public utilities and cooperatives may instead rely more heavily on direct assignment, alternative financing, rate stability objectives, or member/customer governance. These differences imply that similar load shocks may produce larger retail price effects for IOUs than for public utilities or cooperatives. Fifth, electricity market structure could shape both magnitudes and mechanisms.

4 Data

4.1 Data Sources

Data Centers. Our primary data source is Aterio's *US Data Center Locations and Power Demand* product, a facility-level inventory of every active and proposed data center in the United States. As of April 2026, the dataset covers 5,937 facilities: 1,961 active, 735 under construction, and 3,241 announced.[9] For each facility, we observe location (GPS coordinates), operator, building size, estimated power capacity (MW), and activation year. Aterio assembles these records from press releases and company websites, high-resolution satellite imagery, utility and interconnection filings, and permitting records from local and federal agencies. Importantly, every record is manually validated by analysts. Additional details on this dataset are provided in Appendix B.1.

Electricity Prices. We calculate the average retail price as total retail revenue divided by total retail sales, using EIA-861 data at the utility-by-state-year level. The data report total retail sales and a breakdown by customer class (residential, commercial, and industrial). For utilities in restructured states that report both bundled service (in which the utility provides both energy and delivery) and delivery-only service (in which customers buy energy from a competitive supplier), we use the bundled-service price.[10] All dollar variables are deflated to 2024 dollars using the CPI. We exclude utilities that ever file the EIA-861 Short Form, because the short form does not report sales by customer class.[11]

For part of our mechanisms analyses, we construct annual utility-state measures of wholesale electricity prices using day-ahead locational marginal prices (LMPs) published by U.S. regional transmission organizations (RTOs) and independent system operators (ISOs).[12] The data cover

[9] The unit of observation is the physical building, so a multi-tenant colocation facility shared by many operators is counted once. This convention differs from inventories (e.g., DataCenterMap.com) that count at the operator or tenant-deployment level and can list the same building multiple times. The reported number of active U.S. data centers therefore varies across sources: Bloomberg Terminal lists 1,640 as of 2025; the IM3 Open-Source Data Center Atlas lists 1,479 as of February 2026 (Mongird et al., 2026); Muller (2026) reports 2,465 as of 2025 based on S&P data; and DataCenterMap.com lists 4,286 total facilities across 1,837 operators.

[10] An all-in rate for delivery-only customers cannot be constructed from EIA-861 (see Appendix B.1).

[11] The Short Form is filed by utilities with retail sales below 100,000 MWh in the prior year, which are predominantly small municipal utilities and rural cooperatives that account for a small share of total sales.

[12] We obtained wholesale prices through power.dev, which provides prices originating from the individual RTOs/ISOs.

seven organized wholesale markets over 2015–2024. LMPs reflect the marginal cost of supplying electricity at a given location, incorporating energy costs, transmission congestion, and losses. We link these prices to utility-state pairs using utilities' market affiliations and service territories. We construct measures for total LMP and its energy, congestion, and loss components, expressed in 2024 dollars. Appendix B.1 provides the construction details.

Electric Utility Service Territories, Characteristics, and Operational Profile. We gather additional utility-level information from EIA-861, including each utility's service territory (the set of counties it serves), balancing authority, ownership category, and total number of customers along with the breakdown by customer class. We group ownership into three categories: investor-owned utilities (IOUs), cooperatives, and public utilities (including federal, state, municipal, and political subdivision utilities). Utilities' operational profiles come from the EIA-861 operational data, which report summer peak demand, winter peak demand, and total energy disposition at the utility-year level. Because these variables are not reported separately by state, they capture each utility's system-wide operations. We define annual peak demand as the maximum of the utility's summer and winter peaks, measured in megawatts. We then construct the load factor as total energy disposition divided by the product of annual peak demand and the number of hours in the year. This measure captures average system load as a share of peak load: higher values indicate flatter and more continuous demand, and therefore greater utilization of existing capacity.

Infrastructure Investment. We collect annual utility-level infrastructure investment data from S&P Capital IQ Pro, which compiles FERC Form 1 filings. The panel covers 1988–2025 for investor-owned utilities (IOUs).[13] For each utility-year, we construct three groups of outcome measures. The first captures investment in physical plant: gross additions to generators, transmission plant and overhead conductors, and distribution station equipment; electric plant in service; and construction work in progress (CWIP) for electric plant, which records spending on projects not yet in service. The second group captures transmission spending in greater detail: total transmission operation and maintenance (O&M) expense, and expenditures on new and existing transmission lines. The third group captures one of the ways that utilities finance and recover these investments: customer advances for construction (CAC), which are upfront cash contributions customers make toward infrastructure built to serve them. All dollar variables are deflated to 2024 dollars. To avoid double-counting, we drop parent utilities whose filings consolidate subsidiary operations and retain only subsidiary-level reports.[14]

Utility Rate Cases. We use data on resolved electricity rate cases from S&P Capital IQ Pro's Regulatory Research Associates (RRA) database. The data cover 2010–2025 and identify the filing utility, jurisdiction, case type (e.g., vertically integrated, distribution-only), decision type (settled,

[13] Municipal utilities and cooperatives are not required to file Form 1 and are therefore absent.
[14] The holding parent utility companies do not have an EIA code.

fully litigated, or rate-case-by-rate-case), duration, and the dates of initial filing and final decision. We observe utilities' requested values and commissions' authorized values for the change in revenue requirement, the return on equity (ROE), the rate base, and the return on rate base. We restrict the sample to electric service cases and link each utility to our main dataset using the EIA utility code provided by RRA.[15] Because RRA covers only rate-regulated utilities, the resulting panel consists of IOUs. We collapse the case-level data to a utility-state-year panel by each case's filing year.[16]

Large-Load Tariffs. To capture how utilities are now starting to price and recover the cost of serving large new loads, we draw on the Halcyon Large Load Tariff Tracker (March 2026 release), which catalogs special tariffs, contracts, riders, and frameworks that utilities are proposing or adopting for large customers such as data centers. For each utility-state, we build an indicator for whether it ever has any large-load tariff along with indicators for the cost-recovery provisions: a minimum monthly demand (take-or-pay) payment, an upfront contribution in aid of construction (CIAC), a long-term contract or demand requirement, and an exit or early-termination fee. We summarize a tariff's overall stringency with an indicator for a strong cost-recovery tariff equal to one when at least three of the four core provisions are present.

Additional Data. We measure each utility-state's exposure to fossil fuel generation using the generation mix of the balancing authority (BA) in which it operates. We construct BA-level fossil fuel shares using generator- and plant-level data from EIA-923 and EIA-860. We identify each generator's fuel type, net generation, and BA region, and then aggregate net generation to the BA-by-fuel-type level. We define a BA's fossil fuel share as the share of total net generation from fossil fuels, including coal, natural gas, oil, and petroleum coke. We average this measure over 2005–2007, a pre-sample window that does not overlap with the 2008 financial crisis to avoid capturing crisis-related shifts in generation patterns. We then assign each utility-state pair the fossil fuel share of the BA in which it operates.

Reserve margins are constructed from EIA-411 data over 2005–2007, defined as available capacity above peak demand as a percentage of peak demand. They are calculated at the NERC-region level as capacity resources minus net internal demand, divided by net internal demand and multiplied by 100. Lower values therefore indicate tighter, more capacity-constrained grid conditions. We map reserve margins to the balancing-authority level. For balancing authorities that span multiple NERC regions, we use a generation-weighted average of the corresponding NERC-region reserve margins, with weights given by the balancing authority's generation share in each NERC region. Finally, we assign each utility-state pair the reserve margin of the balancing authority in which it operates.

[15] We exclude natural-gas-only cases and combined-utility cases for which the electric portion cannot be separately identified.

[16] Requested and authorized ROE and returns on rate base, as well as authorized rate base, are defined only for utility-state-years with a rate case reporting them.

We construct regional economic characteristics using data from two sources. County-level total and manufacturing gross domestic product (GDP), as well as population, come from the Bureau of Economic Analysis (BEA) Regional Economic Accounts. Counts of total and manufacturing establishments come from the U.S. Census Bureau's County Business Patterns. We average each county-level series over 2005–2007, a pre-sample window that ends before the 2008 financial crisis. We then aggregate these measures to the utility-state level across all counties in the utility-state's service territory. Using these aggregated measures, we construct three regional economic characteristics: total population, GDP per capita, and manufacturing establishments per capita.

4.2 Construction of Final Dataset for Analysis

We aggregate the data center records to the county-year level, computing newly activated capacity (MW) and the count of newly activated facilities in each county and year. We keep county-years with newly activated capacity above 5 MW to exclude facilities that are too small to materially affect the grid. For each remaining county, we take the earliest entry year as the treatment start, and we also construct continuous measures of the cumulative facility count and cumulative capacity in each county at the end of each year.

We match this county-level entry information to utility-state-level electricity prices through utilities' service territories.[17] For each utility-state pair, we take the earliest data center entry year among the counties it serves in that state as the treatment start. The continuous measures—cumulative facility count and cumulative capacity—are constructed for each utility-state-year by summing the corresponding county-year values across all counties the pair serves in that state.

We restrict the sample to the contiguous U.S. and to 2010–2024 for three reasons. First, activation years before 2005 are often imprecise: many early entries are inferred rather than directly observed, and Aterio's satellite-based verification is unavailable for facilities built before high-resolution commercial imagery became widely available. Second, pre-2010 data centers differ substantially from those coming online later. Facilities built during the dot-com and early-enterprise era were mostly internet data centers hosting websites, email, and corporate ERP systems, with typical loads of kilowatts to a few megawatts. The hyperscale era, with single-tenant campuses of hundreds of megawatts operated by cloud and content providers, began around 2010 with the first generation of purpose-built Facebook, Google, and AWS sites. Including pre-2010 facilities would mix two qualitatively different technologies and load profiles. Third, the 2010 start avoids confounding effects from the 2008–2009 financial crisis for both data center investment and electricity demand.

For the rate case data, we match data center entry to utility-state pairs using each utility's state-specific service territory, as in the main retail price analysis. For the infrastructure investment

[17] For example, if a data center enters Durham County, every utility-state pair whose service territory covers Durham County is treated. If a utility serves multiple states but only one sees data center entry, the pairs involving that utility in the non-entry states are controls.

data, FERC Form 1 filings are not reported separately by state, so we match each county-level data center entry to each utility's combined service territory across all states it serves. We restrict the rate case and infrastructure investment samples to 2010–2024 and retain only utilities that appear in the main retail price sample. We merge the large-load tariff data to the retail price panel by utility-state. Utilities absent from the tracker are coded as not having a large-load tariff.

4.3 Descriptive Statistics

Our final main dataset covers 14,733 utility-state-year observations spanning 1,048 utility-state pairs and 973 distinct utility companies operating across 48 states.[18] Table 1 reports summary statistics for the key variables. Variables in Panels A, B, and D are at the utility-state-year level, while those in Panels C and E are at the utility-year level.

The average retail price is \$0.13/kWh across all customer classes, highest for residential customers (\$0.14/kWh) and followed by commercial (\$0.13/kWh) and industrial customers (\$0.10/kWh). The number of observations after data center entry account for about 6% of utility-state-years. On average, the cumulative data center capacity is around 5 MW unconditionally but 80.6 MW among the 908 utility-state-years that host at least one large data center, indicating substantial heterogeneity in exposure.

Panels C and D report summary statistics for variables on infrastructure investment and rate cases, which only include IOUs. Infrastructure spending is dominated by transmission, with \$634 million per utility-year spent on existing lines and \$327 million in construction work in progress. As for rate cases, 27% of utility-state-years involve a rate-case filing, with general cases (23%) far more common than limited-issue rider cases (5%). The average authorized ROE is around 0.7% lower than the average requested ROE, and the average authorized rate base is \$3.36 billion. Finally, operational characteristics in Panel E show an average load factor of 0.56 and average peak demand of 440 MW.

Figure 1 maps cumulative data center capacity by county for facilities activated during 2010-2024, the entry period used in our analysis. The vast majority of counties host little or no capacity, and activity is highly concentrated in a small number of counties. The highest-capacity counties cluster most heavily in northern Virginia, Texas, Arizona, and the Pacific Northwest. Appendix Figure A1 shows the broader footprint of all facilities activated by the end of 2025, including facilities outside the 2010-2024 entry window.

Appendix Figures A2 and A3 plot the expansion of data centers over time, separately by utility ownership and regulatory environment. Investor-owned utilities (IOUs) lead on both margins: both the share of utilities hosting at least one active data center and their average cumulative capacity. Data center expansion proceeds much more slowly for publicly-owned utilities and cooperatives. Across industry structures, average cumulative capacity rises faster in restructured

[18] Only 57 utilities serve customers in more than one state.

than in non-restructured states. While the trends start to diverge as early as 2014, the gap widens markedly after 2019 and then diverges even more toward the end of the sample.

Appendix Table A1 reports descriptive statistics separately for restructured and non-restructured states. Restructured states have higher average retail prices and greater data center exposure on the extensive margin, although cumulative capacity conditional on entry is similar across the two groups. These baseline differences motivate our heterogeneity analysis by electricity industry structure. They also reinforce the importance of the fixed effects and robustness checks described below, which identify effects from within-utility-state changes over time and allow utilities and regions with different baseline profiles to follow different trends.

5 Empirical Strategy

We estimate the causal effect of data center entry on retail electricity prices using variation in the timing of large data center activations across utility service territories. The unit of observation is a utility-state pair by year. Treatment begins in the first year in which a large data center activates in any county served by the utility-state pair. Since treatment is defined as first entry, we drop pairs whose first large data center activation occurred before the start of our sample in 2010, for which we do not observe a pre-period.

A natural starting point is the two-way fixed effects (TWFE) specification:

$$Y_{it} = \beta \, \text{PostDC}_{it} + \alpha_i + \delta_t + \varepsilon_{it}, \tag{12}$$

where Y_{it} is an outcome for utility-state pair i in year t, PostDC_{it} equals one in all years on or after the first data center activation, α_i are utility-state pair fixed effects, and δ_t are year fixed effects. The pair fixed effects absorb time-invariant differences across utility-state pairs (e.g., commission structure, generation mix, climate, baseline industrial composition, persistent rate level disparities, etc.), while year fixed effects absorb common annual shocks such as fuel price movements, macroeconomic conditions, and national trends in electricity demand. Standard errors are clustered at the utility-state-pair level, the level of treatment assignment and outcome measurement.

Since data center entry is staggered over time and treatment effects may evolve after entry, our preferred estimates use the difference-in-difference framework of Callaway and Sant'Anna (2021) (hereafter CSDID). Let G_i denote the first year in which utility-state pair i receives a large data center exposure, with $G_i = \infty$ for never-treated pairs. For each treated cohort g and year $t \geq g$, we estimate the group-time average treatment effect:

$$\text{ATT}(g,t) = \mathbb{E}\left[Y_{it}(g) - Y_{it}(\infty) \,\middle|\, G_i = g \right], \quad t \geq g, \tag{13}$$

where $Y_{it}(g)$ is the potential outcome under first treatment in year g, and $Y_{it}(\infty)$ is the never-treated

potential outcome. Each $\text{ATT}(g,t)$ compares the outcome for cohort g to the corresponding change among never-treated pairs. This avoids comparisons that use already-treated units as controls for later-treated units. We implement the improved doubly robust estimator of Sant'Anna and Zhao (2020), combining outcome regression with inverse-probability weighting.

We summarize the group-time effects in two ways. First, we report event-study estimates by event time $e = t - g$:

$$\text{ATT}(e) \;=\; \sum_{g} w_{g,e} \cdot \text{ATT}(g,\, g+e), \tag{14}$$

where $w_{g,e}$ weights cohorts by their contribution to event time e. These estimates trace the dynamic path of outcomes before and after first data center activation and provide pre-treatment placebo tests. Second, we report an overall ATT:

$$\text{ATT} \;=\; \sum_{g} \sum_{t \geq g} w_{g} \cdot \text{ATT}(g,\, t), \tag{15}$$

which averages post-treatment effects across cohorts and years. Because treatment is defined by first data center entry, these estimates capture the effect of initial entry and any subsequent data center growth in the same utility-state pair.

5.1 Identification Strategy

The main empirical challenge is that data centers do not locate at random. Firms select locations based on electricity prices, land availability, digital connectivity, tax incentives, permitting, and grid conditions. Many of these factors may also be correlated with the outcomes we study. Utility-state pairs that receive data centers may differ from those that do not in system size, baseline prices, customer composition, generation mix, electricity industry structure, local economic growth, and expected future electricity demand. Our empirical strategy therefore does not require data center siting to be random. Instead, it requires that, absent data center entry, treated and comparison utility-state pairs would have followed parallel trends in retail prices and related outcomes. We address the plausibility of this identifying assumption in several ways.

First, the baseline DID design exploits within utility-state pair variation over time. Utility-state-pair fixed effects absorb time-invariant differences across pairs, including persistent differences in commission structure, fuel mix, baseline industrial composition, climate, and rate levels. Year fixed effects absorb common shocks to all utilities, such as national fuel price movements, macroeconomic conditions, common trends in electricity demand, and federal policy changes.

Second, we restrict the analysis sample to improve comparability and measurement. We exclude utilities that file the EIA-861 short form, which are generally smaller utilities with less detailed reporting and different customer composition. We retain only utility-state pairs with a retail component and require pairs to serve residential, commercial, and industrial customers. These

restrictions ensure that the customer classes used to measure incidence are observed consistently across treated and comparison groups.

Third, we assess robustness to differential time trends. A concern is that data centers select into utility-state pairs already on different trajectories because of regional economic growth, electricity market reforms, renewable policies, grid conditions, or differential demand trends. We therefore estimate specifications with utility-specific and grid region-specific trends. We also allow year effects to vary with pre-period utility-state characteristics, including customer base, average retail price, and total sales, as well as pre-period regional characteristics, including manufacturing establishments per capita, population, and GDP per capita. These specifications allow utilities and regions with different baseline profiles to follow different time paths. The estimates remain similar across these alternative specifications.

Fourth, we implement a matched DID design using entropy balancing. This approach re-weights never-treated utility-state pairs so that their pre-period characteristics match those of treated pairs. We balance on pre-period average retail prices, customer counts, manufacturing establishments per capita, and GDP per capita, and then apply the Callaway and Sant'Anna (2021) estimator to the weighted sample. The matched estimates are generally larger than the baseline estimates, suggesting that observable pre-period differences are unlikely to be driving the main results upward.

Parallel Trends Assumption. We use event-study estimates to assess pre-treatment dynamics. The key identifying assumption is that, absent data center entry, treated utility-state pairs would have evolved similarly to never-treated pairs. While this assumption cannot be tested directly, the pre-treatment event-study coefficients are small and statistically indistinguishable from zero across the main outcomes. This provides suggestive evidence against differential pre-trends in the years before the initial data center activation.

SUTVA. A final concern is interference across utility-state pairs (i.e., a SUTVA violation). Data center entry in one service territory could affect other territories through wholesale markets, transmission cost allocation, or multi-state utility cost allocation. Wholesale market spillovers would tend to attenuate our estimates if comparison utilities face some of the same market-level price pressure as treated utilities. Cross-state cost allocation is a more direct concern for multi-state utilities. That is, a utility-state pair without data center entry may be partially treated if the same utility serves another state with data center entry and allocates some costs across its system. Our utility-state design mitigates this concern relative to a utility-level design, and in robustness checks we exclude from the control group utility-state pairs whose utility experiences data center entry in another state.[19] The baseline price estimates are essentially unchanged.

[19] In our sample, only 14 of 112 treated utilities are exposed to this channel.

6 Main Results

6.1 Retail Prices and Customer Class Incidence

We begin by examining the average effect of data center entry on retail electricity prices. Figure 2 presents event study estimates tracing the evolution of electricity rates before and after initial data center activation, and Table 2 reports corresponding average treatment effects. Two main findings emerge. First, retail electricity rates increase following data center entry. Retail prices rise by approximately 2.7% on average, with little evidence of systematic differential pre-trends prior to treatment. The effects grow gradually over time, consistent with both the multi-year processes through which utilities invest in infrastructure and recover costs as well as continued data center expansion following initial entry, as facilities tend to cluster geographically (Fang and Greenstein, 2025).

Second, these effects differ across customer classes. Residential prices increase by 2.1%, commercial prices by 2.8%, and industrial prices by 4.2%, reflecting how average retail rates mask important heterogeneity in how the costs associated with serving new demand are allocated across ratepayers. The positive residential effect indicates that at least some of the costs of data center growth are passed through to households. The larger effects for commercial and (particularly) industrial customers also suggest that non-residential customers might bear a larger share of the price increase, although the pass-through to incumbent commercial and industrial ratepayers should be interpreted with care: depending on utility tariff classifications, data centers themselves may be recorded as commercial, industrial, or special large-load customers. As a result, the commercial and industrial effects may reflect both pass-through to incumbent firms and the prices paid by the new large-load customers generating the demand shock.[20]

Since treatment is defined by the first data center activation, the baseline estimates should be interpreted as the effect of entering into a period of data center growth rather than the effect of a single facility alone. Appendix Table A2 reinforces this interpretation by splitting treated utility-states according to whether they experience additional data center entry after the initial activation. The price effects are much larger and statistically significant among utility-states with subsequent entry: average retail prices rise by 3.3%, residential prices by 2.9%, commercial prices by 4.1%, and industrial prices by 4.4%. On the other hand, estimates are smaller and statistically insignificant among utility-states with no subsequent entry. These patterns are consistent with the interpretation that the main results capture the cumulative effect of data center expansion.

[20]Customer classifications in EIA-861 are based on utilities' reported rate classes rather than a uniform end-use classification. As a result, data centers may appear in commercial, industrial, special contract, or large-load categories depending on the utility's definition. Appendix B.1 provides additional detail.

6.2 Institutional Heterogeneity

We now turn to examining whether the price responses following data center entry differ across electricity sector institutions. As discussed in Section 3, utilities differ in their ownership structures and industry structures, which determine investment incentives, financing, and cost-recovery strategies. The distribution of data center exposure also differs (see Appendix Figures A2 and A3).

6.2.1 Utility Ownership

Investor-owned utilities (IOUs), public utilities, and cooperatives differ in how infrastructure investments are financed, how costs are recovered, and how costs are allocated across customers. We therefore examine whether the retail price effects of data center entry vary systematically across ownership types.

Figure 3 presents ownership-specific event-study estimates, while Table 3 reports the corresponding average treatment effects. The results reveal striking differences. Among IOUs, data center entry increases average retail prices by 5.6%, more than twice the magnitude of the baseline effect in Table 2. The effects are statistically significant across all customer classes for IOUs: residential prices increase by 5.2%, commercial prices by 4.6%, and industrial prices by 8%. The event study estimates show little evidence of differential pre-trends and indicate that these effects emerge gradually over time, consistent with both the time required for infrastructure investment and regulatory cost recovery, and continued data center expansion and clustering following initial entry.

In contrast, the effects are much smaller and less systematic for public utilities and absent for cooperatives. Public utilities exhibit a modest increase in average retail prices (2.4%), but the customer class estimates are generally smaller and statistically indistinguishable from zero. We find no evidence of price responses for cooperatives.

6.2.2 Electricity Industry Structure

We next examine whether the retail price effects differ between utilities operating in states with restructured electricity industries and those in non-restructured states given the potential implications for costs. In non-restructured states, vertically integrated utilities often own or procure generation, transmission, and distribution through integrated planning. In restructured states, generation and delivery are separated, so supply is more often procured through wholesale markets, default service procurement, contracts, or retail suppliers, while distribution utilities recover regulated network costs through rates.

Table 4 reports average treatment effects separately for utilities in restructured and non-restructured states. Overall, retail price effects are substantially larger in restructured markets. Following data center entry, average prices increase by 3.8% in restructured states, compared with

a smaller and statistically insignificant 1.5% increase in non-restructured states. Residential prices increase by 4% in restructured states but do not increase in non-restructured states. Commercial prices increase in both settings, while industrial price effects are larger in restructured states but are not precisely estimated.

Appendix Table A3 examines the two-way split between ownership and regulatory environment. These results show that ownership is the first-order source of heterogeneity: price effects are concentrated among IOUs in both restructured and non-restructured states, while public utilities and cooperatives show little systematic response in either environment. Within IOUs, effects are generally larger in restructured states, especially for average and residential prices, but IOUs in non-restructured states also experience meaningful price increases. This pattern suggests that IOU ownership is central to retail pass-through, while industry structure shapes the magnitude and potentially the mechanisms of pass-through, which we explore in Section 7.

6.3 Magnitude of Effects

To put these results in perspective, we translate the percentage change estimates into electricity bill effects using pre-treatment electricity prices and sales for treated utilities in each estimation sample. The average 2.1% increase in residential rates corresponds to an increase of 0.32 cents per kWh, which is roughly $3.21 per month ($39 per year) for the average residential customer. The magnitudes are larger in the institutional settings where the price effects are concentrated. Among IOUs, where average residential prices are also higher, the 5.2% residential effect corresponds to an increase of about 0.88 cents per kWh, or roughly $7.34 per month ($88 per year) on average. In restructured states, the 4% residential effect corresponds to about $5.94 per month ($71 per year) on average.

These estimates are smaller than some of the bill increases discussed in recent public debates around data centers, but an important caveat to keep in mind is that our estimates are for realized average retail price effects over our sample period (data centers that enter between 2010 and 2024). They are not forward-looking projections, and they do not capture some of the most recent hyperscale developments or high-growth hotspot regions.[21]

At the same time, the incidence is not negligible. Even modest monthly increases can matter for households already facing high energy burdens. In levels, the dollar implications are also much larger for commercial and industrial customers because their electricity consumption is higher. In the full sample, the estimated effects imply average annual bill increases of roughly $225 for commercial customers and about $52,000 for industrial customers, with larger industrial effects among IOUs. We interpret the estimates as evidence of measurable early pass-through from data center growth.

[21] Some utility-states with first data center entry before 2010 are also areas with large or many data centers (e.g., some utility-states in Virginia).

6.4 Distributional Implications

A remaining salient concern is whether price effects are larger in places where ratepayers may be more vulnerable to higher energy costs. Appendix Table A4 splits the sample by pre-treatment average electricity prices and GDP per capita. Panel A shows that, while price effects are similar for the two groups when examining all customers together, class-specific effects are somewhat larger in areas with above-median pre-treatment electricity prices. The differences are modest, though, and should be interpreted cautiously because baseline prices are correlated with utility ownership and regional characteristics. Panel B shows larger effects in lower-GDP areas for average, residential, and industrial prices. While commercial price effects are larger in higher-GDP areas, the residential pattern is especially relevant for incidence. These results suggest that data center entry has generated somewhat larger retail price increases in places that were already facing higher electricity prices or lower levels of economic prosperity, but they do not measure household-level energy burdens or income-specific pass-through directly.

6.5 Robustness Checks

Addressing Confounding Factors. We assess whether the main price results are robust to alternative comparison groups, controls, samples, and inference procedures. First, Panels A–H of Appendix Table A5 report results from when allowing treated and comparison utility-state pairs to follow different time paths. The estimates remain similar when adding utility-specific trends, grid-region-specific trends, interactions between year fixed effects and pre-period utility characteristics, and interactions between year fixed effects and pre-period regional characteristics.[22] These specifications address concerns that data centers locate in utilities or regions already on different price trajectories because of baseline prices, scale, industrial composition, income, industry structure, or regional grid conditions.

Matched DID Design. We also implement a matched DID design using entropy balancing (Hainmueller, 2012).[23] Treated utility-state pairs differ from never-treated pairs in the raw sample—before entry, they are larger, have higher retail prices, and operate in higher-income areas. We balance on four pre-period characteristics measured from 2005-2007: average retail price, number of customers served, manufacturing establishments per capita, and GDP per capita. Appendix Table A6 shows that entropy balancing removes these observable pre-period differences, and Appendix Figure A4 and Appendix Table A7 show that the matched estimates remain positive and are generally larger than the baseline estimates. This pattern suggests that observable pre-period imbalances are unlikely to be driving the main results upward.

[22] See Appendix B.2 for details on how we assign utility-state pairs to grid regions.
[23] Additional details can be found in Appendix B.3.

Instrumental Variable Estimates. As a complementary approach, we estimate instrumental variable (IV) specifications of the effects of cumulative data center *capacity*. These specifications differ relative to our baseline design in two key ways: they use the log cumulative data center capacity as the endogenous variable rather than an indicator for first entry, and the sample retains utility-state pairs that already hosted data centers at the beginning of the sample. We construct two IVs that interact historical local exposure to the data-processing and hosting sector in a pre-treatment period (2005–2007)—measured either by establishments per capita or by log employment—with total annual capital expenditures by Google/Alphabet, Amazon, Meta, and Microsoft. Our IV strategy thus identifies the effect of cumulative capacity among utility-state pairs whose data center growth is shifted by historical data-processing exposure. The intuition is that places with pre-existing data-processing activity may be more likely to attract additional data centers when large cloud computing firms expand investment.

The first-stage results are strong, and we find positive and statistically significant estimates for the effect of cumulative data center capacity on retail rates across all customer classes (see Appendix Table A9).[24] Using the establishment-density instrument, the coefficient on log cumulative capacity is 0.073 for average retail prices. Using the employment instrument, the corresponding estimate is 0.078. These results reinforce our conclusion that, on average, data center expansion puts upward pressure on retail rates.

Excluding Early or Late Treated Units. We test whether the estimates are sensitive to the inclusion of treated units observed only briefly. Panel I of Table A5 shows that the estimates are similar when including only treated units for which the first data center entries were between 2012 and 2022, ensuring that each treated unit is observed for at least two years before and after entry. We also address a related concern in the heterogeneity analysis. Figures A2 and A3 show that cumulative data center capacity remains negligible for cooperatives, public utilities, and utilities in non-restructured states until the mid-2010s. Table A10 therefore drops treated utility-state pairs with first entry before 2016 and re-estimates the ownership and industry structure splits. The results are similar.

Addressing Spillovers. We also test whether regional market spillovers or cross-state cost allocation contaminates the control group. Utilities spanning multiple states may recover data center-related infrastructure costs across all customers. If anything, such spillovers should raise prices in untreated utility-states and thus attenuate our estimates, but Table A5 probes this concern further by excluding multi-state utilities with data center entry in only a subset of service states (Panel J), and multi-state utilities altogether (Panel K), identifying the effect solely from single-state utilities. The results are unchanged.

[24] Appendix Table A8 reports strong first stages, with Kleibergen–Paap F-statistics of 48 and 125.

Alternative Standard Error Clustering. We assess the sensitivity of inference to alternative clustering. Panels L and M of Table A5 cluster by state and utility rather than by utility-state. The standard errors remain similar to our baseline.

Ruling Out Alternative Explanations. A final concern is that data center expansion through our sample period coincided with other changes in electricity demand and infrastructure investment—such as electrification, renewable generation and storage expansion, grid modernization, and wildfire response and prevention—which could contribute to retail rate increases (as well as the infrastructure investment responses shown in Section 7).

Table A11 reports results from several robustness checks addressing these alternative explanations.[25] Panels A and B control for installed renewable energy (solar and wind) and storage capacity, respectively, in utilities' service counties. These measures account for differences in the scale of other energy infrastructure. Panel C allows retail prices to follow a linear trend that varies with pre-sample wildfire exposure. Panel D excludes observations in states with high battery electric vehicle stocks. Panel E controls for the share of occupied housing units using electricity as their main heating fuel. These checks examine whether broader electrification accounts for the estimated price effects. Panel F controls for smart meter deployment, measured by advanced metering infrastructure meters per distribution customer to capture grid modernization. Finally, Panels G and H exclude observations in California and in Texas and Virginia, respectively, to assess whether the findings are driven by particular states with distinct policy and electricity market environments. Across these specifications, the estimated effects remain positive and statistically significant for all four price outcomes, with magnitudes close to the baseline.

7 Mechanisms

We now investigate the mechanisms underlying our main results, organized around three related questions that map to the framework in Section 3. First, do utilities respond to data center-induced costs through infrastructure investment and improved utilization? Second, through which regulatory and tariff mechanisms are costs recovered, and do cost-causation institutions attenuate pass-through? Third, to what extent do system constraints and wholesale market exposure contribute to price increases? We focus on IOUs for these analyses given how those are the utilities for which prices increase and also the utilities for which we have comprehensive data on investment and rate cases.

[25] Appendix B.2 describes the data sources, variable construction, and samples used in these checks.

7.1 Infrastructure and Operational Responses

7.1.1 Infrastructure Investments

Figures 4 and 5 present event studies for the main infrastructure outcomes among IOUs. Appendix Figures A6 and A7 present separate estimates for utilities operating in restructured and non-restructured states, and Appendix Table A12 reports the average treatment effects.[26]

The first takeaway is that data center entry induces electricity-specific capital activity. Electric construction work in progress (CWIP), which captures capital projects under construction but not yet fully placed in service, increases following data center entry. This response is not present for common or other CWIP, as shown in Appendix Figure A5, suggesting that the response is specific to electric utility infrastructure rather than a broad expansion in utility activity.[27]

As for completed investment, the clearest result is for generation. On average, data center entry increases generator additions by about $64 million, or roughly 428% of the pre-treatment mean. This effect is statistically significant in both restructured and non-restructured states, with magnitudes of roughly 386% and 458% of the corresponding pre-treatment means, respectively. Thus, generation-related capital investment increases following data center entry in both institutional settings, despite differences in the extent of utility generation ownership.[28]

On the other hand, other infrastructure investments following data center entry differ more across industry structures. Electric CWIP rises by about $187 million in restructured states, or 53% of the pre-treatment mean, but we find no corresponding increase in non-restructured states. Distribution station additions also increase in restructured states, by about $12 million (35% of the pre-treatment mean), while the non-restructured estimate is close to zero. These results suggest that utilities in restructured states exhibit a clearer construction-stage and distribution-equipment response to data center entry.

Transmission-related outcomes also provide suggestive evidence of upgrades. While transmission plant additions and conductor additions remain relatively flat, transmission O&M appears to increase on average (Panel E of Figure 4), and the event studies in Figure 5 also suggest that expenditures in existing transmission lines increase—particularly in non-restructured states—as opposed to new transmission lines. Existing-line expenditures rise by roughly 22% relative to the pre-treatment mean for IOUs, although the estimates are not statistically significant. The near- and medium-run transmission response is thus more likely to involve maintenance, reinforcement, or upgrading of existing assets rather than expansion, but the estimates should be interpreted cautiously.

[26] For these utility-level analyses, we classify a utility as restructured only if all states in its service territory are restructured. Utilities serving at least one non-restructured state, including those operating in both types of states, are assigned to the non-restructured group.

[27] Common CWIP captures shared utility assets such as administrative facilities, information systems, and general plant, while other CWIP captures construction outside the electric system.

[28] Although restructuring generally separated generation from delivery, some utilities in restructured states retained generation assets.

Taken together, the infrastructure results show that data center entry generates real investment responses among IOUs, but the composition of those responses differs across electricity industry structures. Generator investment rises in both restructured and non-restructured states, indicating that utilities make substantial supply-side adjustments to accommodate new load. Restructured state utilities, however, show the clearest increases in construction-stage capital activity and distribution station equipment. Utilities in non-restructured states show significant generator additions and positive, though less precise, transmission O&M and existing line investments. These differences suggest that the key distinction is not whether utilities invest, but how investment is timed, in which assets they invest, and how the resulting costs are absorbed or recovered.

7.1.2 Operational Responses

We next examine peak demand and load factor (see Figure 6). Peak demand captures the maximum load the system must serve and therefore speaks to capacity and reliability needs. Load factor, defined as average load relative to peak demand, captures system utilization. These measures help distinguish between two channels: data centers may raise capacity pressure if they increase peak demand, but they may also attenuate average cost increases if they raise average load more than peak load and allow fixed costs to be spread over more kilowatt-hours.

The split by electricity industry structure reveals different adjustment margins. In restructured states, load factors do not increase meaningfully after data center entry. Combined with the larger CWIP response documented above, this suggests that data center load generates construction-stage cost pressure without a comparable utilization effect. Appendix Figure A9, showing that additions to electric plants in service increase for utilities in restructured states but not non-restructured states, reinforces that capital responses appear earlier in restructured states. In non-restructured states, by contrast, load factors rise several years after entry even as peak demand also trends upward, indicating that average load rises more than peak load. This pattern is consistent with utilities partially absorbing new load through higher utilization and fixed-cost spreading.

Overall, these results help explain why infrastructure responses in both market structures do not produce the same retail price effects. Utilities in non-restructured states appear to benefit from a utilization offset that muted near-term price pressure. Utilities in restructured states do not exhibit the same load-factor improvement, making construction, procurement, capacity, and delivery-related costs more likely to translate into retail rate pressure. We next examine how these costs enter the regulatory cost-recovery process.

7.2 Cost Recovery Strategies

The infrastructure and operational results show that IOUs respond to data center entry, but that investment responses need not translate into similar retail price effects. We next examine the regulatory and tariff mechanisms through which costs are recovered. These mechanisms determine

whether new investment enters the revenue requirement, how quickly costs are recovered, and whether costs are socialized across ratepayers.

7.2.1 Rate Cases and Authorized Cost Recovery

Table 5 examines rate case activity. Across all IOUs, data center entry does not significantly affect rate case activity on average, but this masks substantial heterogeneity by electricity industry structure. Among utilities operating in restructured states, data center entry increases the probability of filing any rate case, the probability of filing a general rate case, the probability of filing a rider case, and the total number of cases. Case duration also declines. These patterns suggest a more active and faster cost-recovery process in restructured states, with rider proceedings providing a targeted channel for recovering specific infrastructure or system costs. By contrast, we find no comparable increase in rate case activity among IOUs in non-restructured states.

Table 6 examines what utilities request and what regulators authorize. The clearest result is that, among rate case proceedings with reported rate base values, data center entry is associated with larger authorized rate bases (in both restructured and non-restructured states). This finding indicates that regulators recognize additional capital as prudent, used and useful, and eligible for cost recovery, linking the investment evidence above to the regulatory process for recovering costs.

At the same time, we do not find higher authorized revenue requirements. Requested revenue requirements increase in restructured states, suggesting that utilities perceive data center entry as creating additional revenue needs, but authorized revenue requirements do not increase proportionally. We also do not find evidence that data center entry increases authorized *returns* on equity or rate base. The price effects therefore do not appear to operate through regulators granting utilities higher returns per dollar of capital, but rather, a larger approved capital base, more active rate-case and rider proceedings, and faster cost-recovery channels.

This distinction helps reconcile the infrastructure and price results. In non-restructured states, data center entry increases authorized rate base but does not generate comparable increases in rate case activity, requested revenue requirements, or retail prices. This is consistent with the utilization improvements that we found for these utilities, as improved load factors and higher sales can mute contemporaneous price effects even when capital investment and rate base growth occur. On the other hand, in restructured states, rate base growth occurs alongside more active proceedings, more rider filings, higher requested revenue requirements, and larger retail price effects, consistent with faster pass-through of investment and system costs when there is less room for improving utilization.

7.2.2 Direct Cost Assignment

We next explore whether utilities assign some infrastructure costs directly to the large-load customers that create them. Appendix Table A15 and Appendix Figure A8 report effects on customer

advances for construction (CAC), which are upfront payments from customers toward infrastructure built to serve them. Higher customer advances could imply that utilities recover at least some data center-related infrastructure costs from the customers inducing the investment.

We find that data center entry increases customer advances for construction on average, with the effect concentrated among utilities operating in restructured states. This suggests that these utilities attempt to assign some infrastructure costs directly to large-load customers. However, because these same utilities also experience retail price increases, customer advances do not appear to fully eliminate pass-through to incumbent ratepayers.

Finally, Appendix Tables A16 and A17 provide suggestive descriptive insights on large-load tariffs. These results should be interpreted very cautiously because the tariff data are available only in recent years and adoption is not randomly assigned. Nonetheless, we find that utilities that *eventually* implement large-load tariffs at some point in later years do not exhibit the same retail price increases as utilities that never implement such tariffs. This pattern is consistent with the idea that provisions such as minimum demand commitments, long-term contract requirements, and infrastructure contributions may attenuate pass-through to other ratepayers.

7.3 System Constraints

We found that data center entry induced infrastructure and cost-recovery responses, but retail price effects are not uniform across industry structures. One potential explanation is baseline system slack. When systems have spare capacity, new load may be absorbed through existing or planned capacity, raising utilization and spreading fixed costs over more kilowatt-hours. When systems are already tight, new load is more likely to require costly procurement, reliability expenditures, capacity-related payments, or accelerated infrastructure investment.

We explore this channel using two pre-period measures of system conditions: reserve margins and fossil generation shares. Reserve margins capture available capacity above peak demand, with lower values indicating tighter systems and less spare capacity. Fossil shares provide a complementary proxy for exposure to thermal generation and fuel-cost-intensive marginal supply, capturing the costliness of serving additional load.

Table A13 provides evidence that system constraints matter. In the full sample, price effects are larger in low-reserve-margin regions, except for industrial customers for whom the estimates are less precise. The pattern is especially pronounced in restructured states: data center entry has little effect in high-reserve-margin regions, but increases average, residential, and commercial prices substantially in low-reserve-margin regions. These estimates are consistent with a capacity-adequacy mechanism. When systems have limited spare capacity, large new loads are less likely to be absorbed through utilization gains and more likely to trigger procurement, reliability, or infrastructure costs that are passed through to retail customers.

Table A14 provides complementary evidence. In the full sample, price effects are not sharply

different across high- and low-fossil regions, suggesting that fossil exposure is not a dominant mechanism. However, within restructured states, effects are generally larger in fossil-heavy regions. In non-restructured states, both the fossil share and reserve margin splits reveal less systematic heterogeneity.

These results help explain why infrastructure responses need not generate the same retail price effects across market structures. Utilities in non-restructured states may have absorbed data center load through owned generation, integrated planning, long-term procurement, high utilization, and gradual cost recovery. This interpretation is also consistent with our previous finding of improved utilization in non-restructured states. In restructured states, where we do not observe the same utilization offset, price effects are concentrated precisely where baseline system conditions suggest less spare capacity. These results reinforce the broader mechanism evidence that data center load raises retail rates most where new demand is difficult to absorb and where institutions transmit the resulting cost pressure to retail customers.

7.4 Wholesale Market Exposure

A final potential channel is wholesale market exposure. Data center load may affect wholesale costs by increasing the marginal cost of energy, worsening congestion, increasing marginal losses, raising capacity obligations, or increasing transmission charges. This channel is particularly relevant for interpreting the larger retail price effects in restructured states, where generation is separated from transmission and distribution, and electricity supply is often procured through wholesale markets, contracts, default service arrangements, or retail suppliers.[29] At the same time, wholesale cost shocks do not necessarily pass-through mechanically to average retail rates.

We first examine wholesale prices directly. Appendix Figure A11 plots event study estimates for total day-ahead locational marginal prices (LMPs) and their energy, congestion, and loss components. The results suggest that entry does not increase the energy component of LMPs. Instead, the congestion and loss components increase, consistent with data center load creating localized network stress rather than a market-wide increase in the marginal cost of generation.

We then test whether wholesale price changes explain the retail price effects by adding wholesale price controls to our baseline specification (see Appendix Table A18). Since wholesale price data are available only for utility-state pairs matched to organized markets, we first re-estimate our baseline model for the matched sample (without wholesale controls) in Column 1.[30] Column 2 then controls for total day-ahead LMPs. On average, the estimated retail price effect is slightly attenuated, falling from 3.8% without wholesale controls to 3.4% with LMP controls. In restructured states, the estimate falls from 4.2% to 3.6%, and in non-restructured states, the estimate remains small and statistically insignificant across specifications. These results suggest that wholesale market exposure explains part, but certainly not all, of the retail price response.

[29] Restructuring should not be conflated with organized wholesale market participation, though.
[30] See Appendix B.1 for details on the wholesale price data.

7.5 Summary of Mechanisms Results

Taken together, the evidence points to data center entry inducing infrastructure responses among IOUs, including electricity-specific construction activity, generation investment, distribution equipment (in restructured states), and larger authorized rate bases. But these investment responses do not translate mechanically into higher retail rates. In non-restructured states, utilities appear better able to absorb new load with more gradual cost recovery, whether that's due to owning generation, integrated planning, or improved utilization. On the other hand, in restructured states, we find earlier construction-stage capital activity, more active rate case and rider proceedings, higher requested revenue requirements, increased customer advances for construction, and stronger price effects in constrained systems. Wholesale market evidence provides additional support, with increases in the congestion and loss components of LMPs. Controlling for wholesale prices slightly attenuates—but does not eliminate—the retail price effects. Retail incidence is therefore greatest where new load is costly to absorb, where utilization gains are limited, and where IOU regulatory and market institutions transmit costs to retail customers.

8 Conclusion

The rapid expansion of data centers is creating substantial new electricity demand, often concentrated in a small number of locations. We study who pays for the complementary infrastructure that this growth requires by linking data center entry from 2010–2024 to utility-level retail prices, infrastructure investment, operational outcomes, wholesale prices, and regulatory proceedings.

Data center entry raises average retail electricity prices, including residential rates. These effects are concentrated among IOUs, are much smaller or absent among public utilities and cooperatives, and are larger in states with restructured electricity industries. We provide additional evidence on the underlying mechanisms. Data center entry induces electric construction, generation investment, distribution equipment in restructured states, and larger authorized rate bases, but price effects depend on whether utilities can absorb the load. In non-restructured states, higher utilization and owned generation appear to mute the price effects. In restructured states, faster investment, stronger cost-recovery responses, and tighter constraints are associated with larger retail price increases.

For policymakers, the relevant question is not simply how much electricity a data center uses, but whether the local system can absorb the load at low incremental cost and whether any required costs are assigned to the customer causing them or socialized across incumbent ratepayers. Future effects could be larger in tighter systems but smaller where high-load-factor customers improve utilization or large-load tariffs directly assign incremental costs. More broadly, the distributional consequences of innovation depend on the institutions governing the complementary infrastructure it requires.

References

Allen, T., and C. Arkolakis. 2022. "The Welfare Effects of Transportation Infrastructure Improvements." *The Review of Economic Studies*, 89(6): 2911–2957.

Alvarez, Fernando, David Argente, Joyce Chow, and Diana Van Patten. 2026. "Data Centers and Local Economies in the Age of AI: A Shift-Share Approach." *NBER Working Paper 35194*.

Atkinson, Scott E., and Robert Halvorsen. 1986. "The Relative Efficiency of Public and Private Firms in a Regulated Environment: The Case of U.S. Electric Utilities." *Journal of Public Economics*, 29(3): 281–294.

Averch, Harvey, and Leland Johnson. 1962. "Behavior of the firm under regulatory constraint." *The American Economic Review*, 52(5): 1052–1069.

Bloomberg. 2026. "AI Buildout Drives 76% Power Bill Jump on Largest US Grid." *Bloomberg*. Updated May 14, 2026. Accessed October 12, 2026.

Callaway, Brantly, and Pedro H C Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

Cicala, Steve. 2015. "When Does Regulation Distort Costs? Lessons from Fuel Procurement in U.S. Electricity Generation." *The American Economic Review*, 105(1): 411–444.

Cicala, Steve. 2022. "Imperfect Markets versus Imperfect Regulation in US Electricity Generation." *American Economic Review*, 112(2): 409–441.

Cicala, Steve. 2025. "Restructuring the Rate Base." *AEA Papers and Proceedings*, 115.

Davis, L., and C. Hausman. 2016. "Market Impacts of a Nuclear Power Plant Closure." *American Economic Journal: Applied Economics*, 8(2): 92–122.

Davis, L., and C. Wolfram. 2012. "Deregulation, Consolidation, and Efficiency: Evidence from US Nuclear Power." *American Economic Journal: Applied Economics*, 4(4): 194–225.

Donaldson, D. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *The American Economic Review*, 108(4): 899–934.

Duggan, Mark, Atul Gupta, Emilie Jackson, and Zachary S. Templeton. 2023. "The Impact of Privatization: Evidence from the Hospital Sector." National Bureau of Economic Research Working Paper 30824.

Duggan, Mark G. 2000. "Hospital Ownership and Public Medical Spending." *The Quarterly Journal of Economics*, 115(4): 1343–1373.

Duggan, Mark, Jonathan Gruber, and Boris Vabson. 2018. "The Consequences of Health Care Privatization: Evidence from Medicare Advantage Exits." *American Economic Journal: Economic Policy*, 10(1): 153–186.

Dunkle Werner, Karl, and Stephen Jarvis. 2026. "Rate of Return revisited." *Journal of Public Economics*, 259.

Elhai, C., and Y. Zhou. 2026. "Data Centers: Location Choice and Grid Externalities." *Working Paper*.

Fabra, Natalia, and Mar Reguant. 2014. "Pass-Through of Emissions Costs in Electricity Markets." *The American Economic Review*, 104(9): 2872–2899.

Fang, Tommy Pan, and Shane Greenstein. 2025. "Where the Cloud Rests: The Economic Geography of Data Centers." *Strategy Science*, 10(4): 281–444.

Färe, Rolf, Shawna Grosskopf, and James Logan. 1985. "The Relative Performance of Publicly-Owned and Privately-Owned Electric Utilities." *Journal of Public Economics*, 26(1): 89–106.

Ganapati, S., J. Shapiro, and R. Walker. 2020. "Energy Cost Pass-Through in US Manufacturing: Estimates and Implications for Carbon Taxes." *American Economic Journal: Applied Economics*, 12(2): 303–342.

Gowrisankaran, G., A. Langer, and M. Reguant. 2026. "Energy Transitions in Regulated Markets." *The American Economic Review*, 116(8): 2928–2961.

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis: An Annual Publication of the Methodology Section of the American Political Science Association*, 20(1): 25–46.

Hart, Oliver, Andrei Shleifer, and Robert W. Vishny. 1997. "The Proper Scope of Government: Theory and an Application to Prisons." *The Quarterly Journal of Economics*, 112(4): 1127–1161.

Hausman, Catherine. 2025. "Power flows: Transmission lines, allocative efficiency, and corporate profits." *American Economic Review*, 115(8): 2574–2615.

Hernandez-Cortes, Danae, Kyle Meng, and Paige Weber. 2026. "The Environmental Costs and Geography of U.S. Data Center Expansion." UC Berkeley Haas Working Paper WP-363.

IEA. 2025. "Energy and AI: World Energy Outlook Special Report." International Energy Agency.

Joskow, Paul. 1974. "Inflation and Environmental Concern: Structural Change in the Process of Public Utility Price Regulation." *The Journal of Law and Economics*, 17(2): 291–327.

Joskow, Paul, and Nancy Rose. 1989. "The Effects of Economic Regulation." *Handbook of Industrial Organization*.

Joskow, Paul L., and Richard Schmalensee. 2025. "Cost-of-Service Regulation of Electricity Distribution Services in the US." In *Handbook on Electricity Regulation*. , ed. Jean-Michel Glachant, Paul L. Joskow and Michael G. Pollitt, 44–67. Cheltenham, UK:Edward Elgar Publishing.

Kay, Owen, Robert Reaser, and Reid Taylor. 2026. "Processing Power: The Effect of Data Centers on Wholesasle Electricity Markets." Federal Reserve Bank of Dallas.

Knittel, Chris, Juan Ramon Senga, and Shen Wang. 2025. "Flexible Data Centers and the Grid: Lower Costs, Higher Emissions?" NBER Working Paper 34065.

Kwoka, John E. 2005. "The Comparative Advantage of Public Ownership: Evidence from U.S. Electric Utilities." *Canadian Journal of Economics*, 38(2): 622–640.

MacKay, A., and I. Mercadal. 2026. "Do Markets Reduce Prices? Evidence from the U.S. Electricity Sector." *Working Paper*.

Mamkhezri, J., X. Sun, and Y. Yang. 2025. "The Hidden Cost of the Cloud: Data Centers and Transmission-Induced Electricity Prices." *USAEE Working Paper No. 25-660*.

Masanet, Eric, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. 2020. "Recalibrating global data center energy-use estimates." *Science*, 367(6481): 984–986.

Mongird, Kendall, Chris Vernon, Casey Burleyson, Kerem Ziya Akdemir, and Jennie Rice. 2026. "IM3 Open Source Data Center Atlas."

Muller, Nicholas Z. 2026. "Measuring the impact of data centers in the united states economy: monetary damage from air pollution and greenhouse gas emissions." 35100, Cambridge:National Bureau of Economic Research.

NCSL. 2026. "As Data Centers Surge, States Weigh Jobs, Power and Local Control." *NCSL News*. Accessed October 8, 2026.

NYTimes. 2026. "The Most Bipartisan Issue Since Beer: Opposition to Data Centers." *The New York Times*. Accessed October 8, 2026.

Reuters. 2026. "Big Tech Data Centers Are Driving Up Power Bills at America's Rust Belt Factories." *Reuters*. Updated July 8, 2026. Accessed October 12, 2026.

Sant'Anna, Pedro H C, and Jun Zhao. 2020. "Doubly robust difference-in-differences estimators." *Journal of Econometrics*, 219(1): 101–122.

Satchwell, Andrew, Natalie Frick, Peter Cappers, Sanem Sergici, Ryan Hledik, Goksin Kavlak, and Glenda Oskar. 2025. "Electricity Rate Designs for Large Loads: Evolving Practices and Opportunities." ark:/13030/qt19m1k8vr, None, 2502077.

Scalera, C., V. Bosetti, and F. Pecci. 2026. "Large Load Electricity Customers and Spillovers on Household Prices: Evidence from U.S. Data Centers." *CEPR Discussion Paper Series DP21854*.

Shehabi, Arman, Sarah J Smith, Alex Hubbard, Alex Newkirk, Nuoa Lei, Md Abu, Bakar Siddik, Billie Holecek, Jonathan Koomey, Eric Masanet, and Dale Sartor. 2024. "2024 United States Data Center Energy Usage Report." Berkeley:Lawrence Berkeley National Laboratory.

Shleifer, Andrei. 1998. "State versus Private Ownership." *Journal of Economic Perspectives*, 12(4): 133–150.

Smith, S., A. Hubbard, A. Newkirk, M. Ganeshalingam, B. Holecek, D. Sartor, M. Mills, and A. Shehabi. 2026. "United States Data Center Energy Usage Report: 2025 Update." Berkeley:Lawrence Berkeley National Laboratory.

WashingtonPost. 2026. "Data Centers Have United Americans of Both Parties in a Shared Hatred." *The Washington Post*. Accessed October 12, 2026.

Watten, A., J. Bistline, and G. Blanford. 2026. "Have Data Centers Raised Your Electric Bill? Causal Evidence from the United States." *Working Paper*.

Wiser, Ryan, Eric O'Shaughnessy, Galen Barbose, Peter Cappers, and Will Gorman. 2025. "Factors Influencing Recent Trends in Retail Electricity Prices in the United States." *The Electricity Journal*, 38: 107516.

Wolfram, C. 1999. "Measuring Duopoly Power in the British Electricity Spot Market." *The American Economic Review*, 89(4): 805–825.

Yue, Daniel, and Yiyang Zeng. 2026. "The Local Economic Effects of Data Center Entry." *Working Paper*.

Main Text Figures

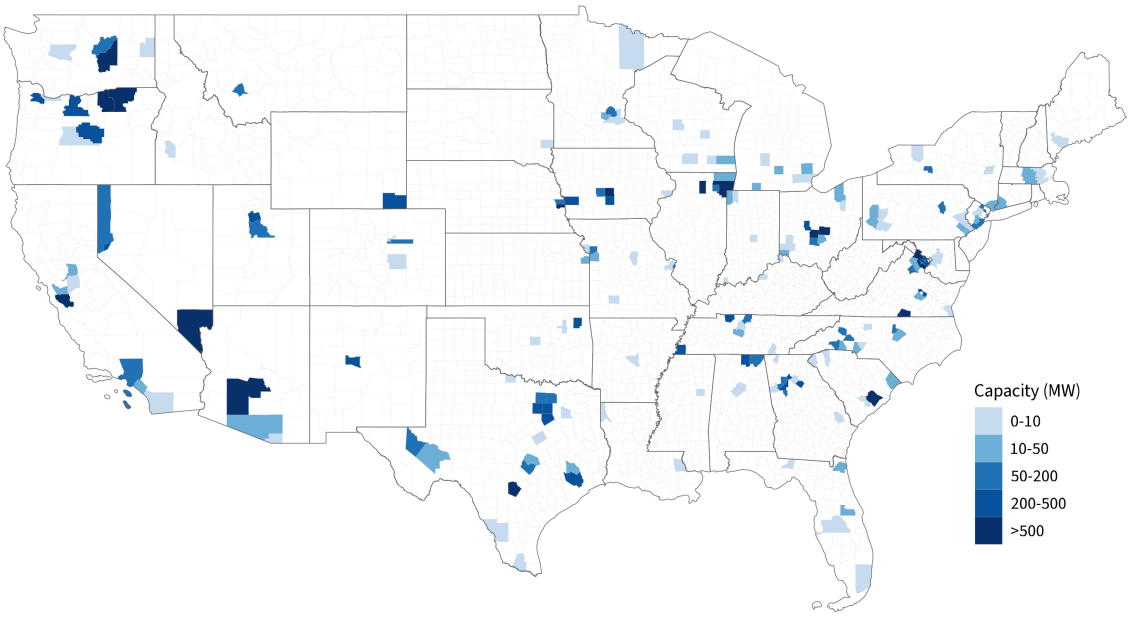


Figure 1: Data Center Capacity by County for Facilities Activated in 2010-2024

Notes: This figure shows total data center capacity (MW) by county in the contiguous United States. It includes facilities activated during 2010–2024, the entry period covered by our analyses.

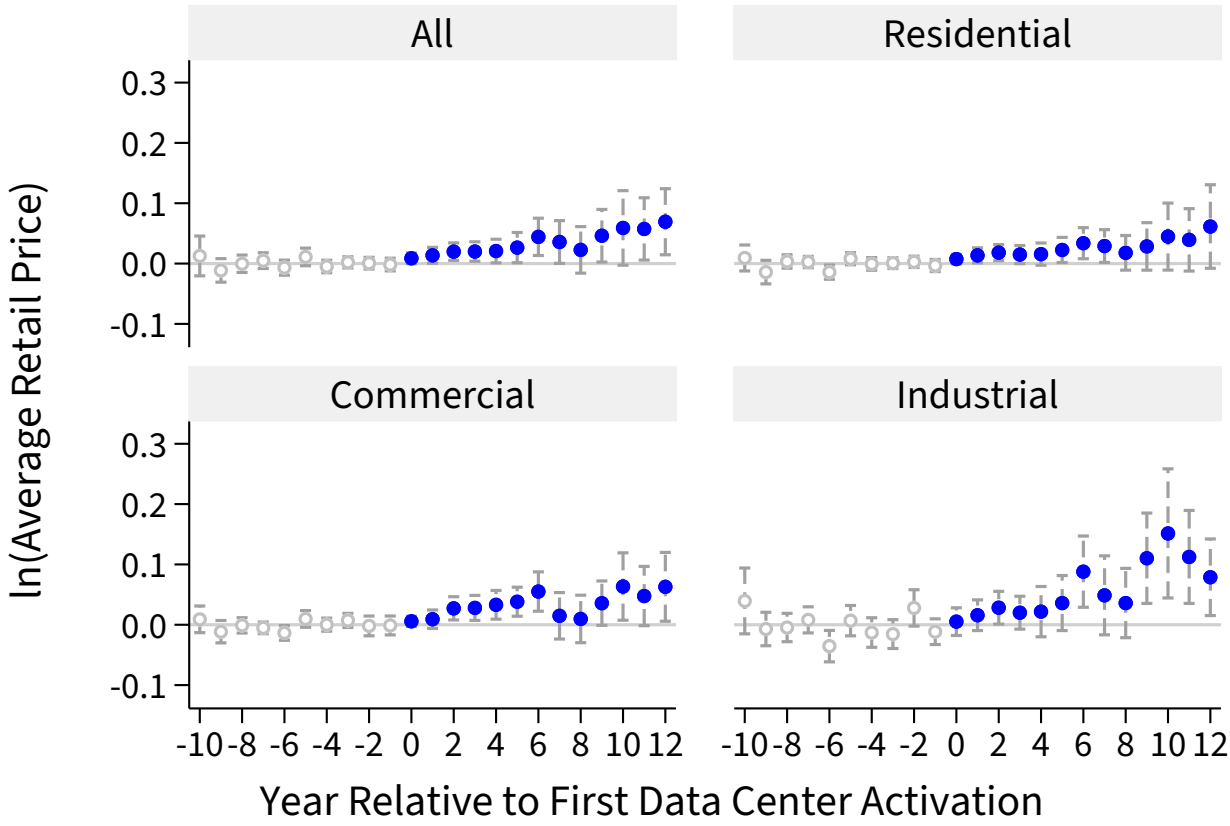
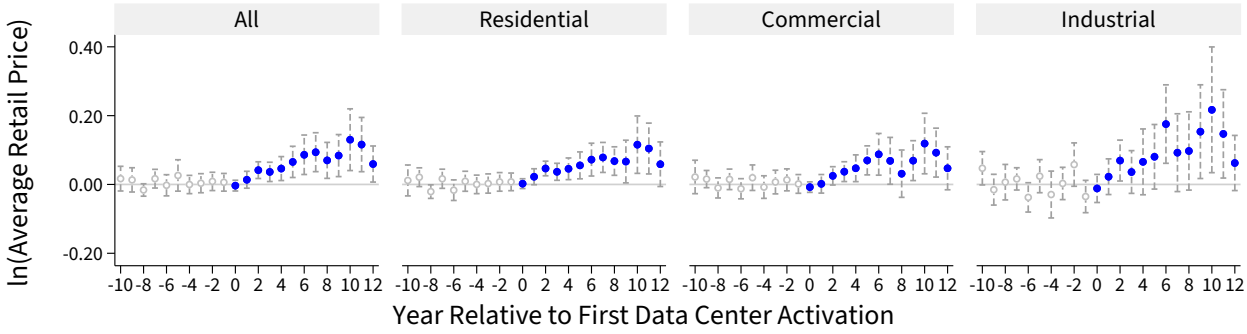
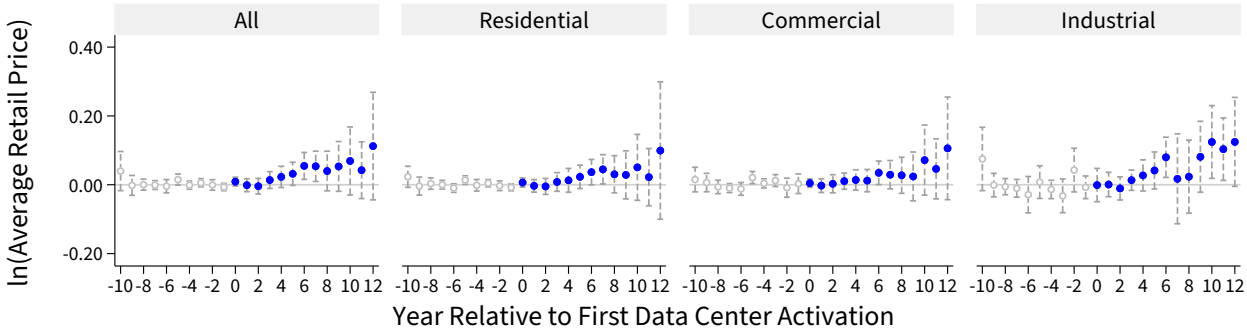


Figure 2: Dynamic Impact of Data Center Entry on Average Retail Prices

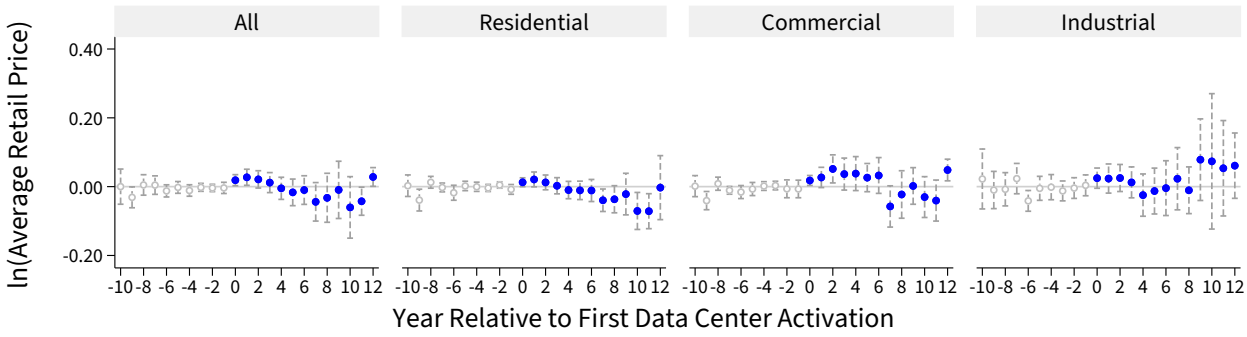
Notes: The unit of observation is the utility-state-year. Figures plot event-study coefficients and 95% confidence intervals, capturing the difference in (logged) average retail electricity rates for utility-state pairs with and without data centers before and after initial data center entry. The first panel shows estimates for average retail rates across all customer classes, whereas the following three panels examine average rates for residential, commercial, and industrial customers separately. Standard errors are clustered at the utility-state level.



(a) Investor-Owned Utilities



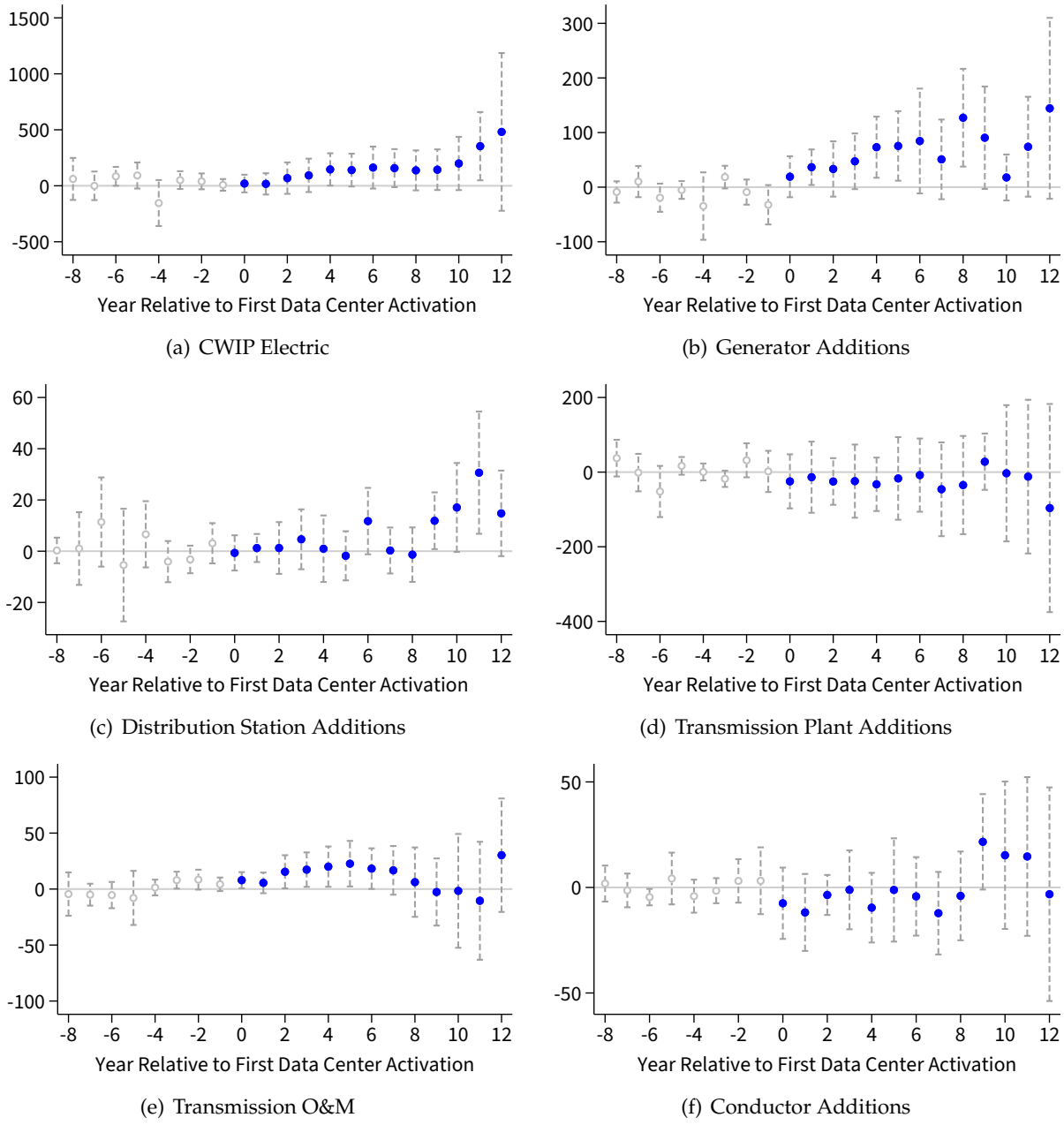
(b) Publicly-Owned Utilities



(c) Cooperatives

Figure 3: Dynamic Impact of Data Center Entry on Average Retail Prices by Utility Ownership

Notes: The unit of observation is the utility-state-year. Figures plot event-study coefficients and 95% confidence intervals by utility ownership, capturing the difference in (logged) average retail electricity rates for utility-state pairs with and without data centers before and after initial data center entry. Panel A shows the effects for IOUs, Panel B for public utilities, and Panel C for cooperatives. For each subsample analysis, we include all utilities without data centers in the control group (despite ownership type) while the treatment group in each case includes only utilities with each ownership type that ultimately have data centers. In each panel, the first column shows estimates for average retail rates across all customer classes, whereas the following three columns examine average rates for residential, commercial, and industrial customers separately. Standard errors are clustered at the utility-state level.



(a) CWIP Electric

(b) Generator Additions

(c) Distribution Station Additions

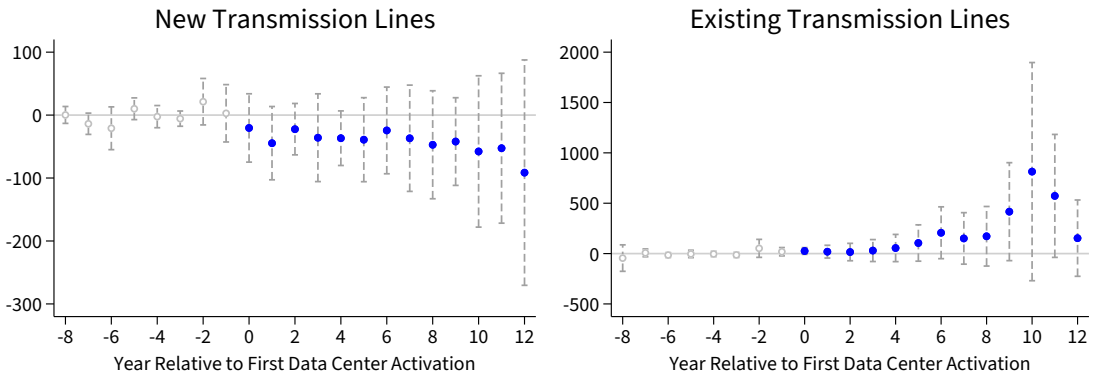
(d) Transmission Plant Additions

(e) Transmission O&M

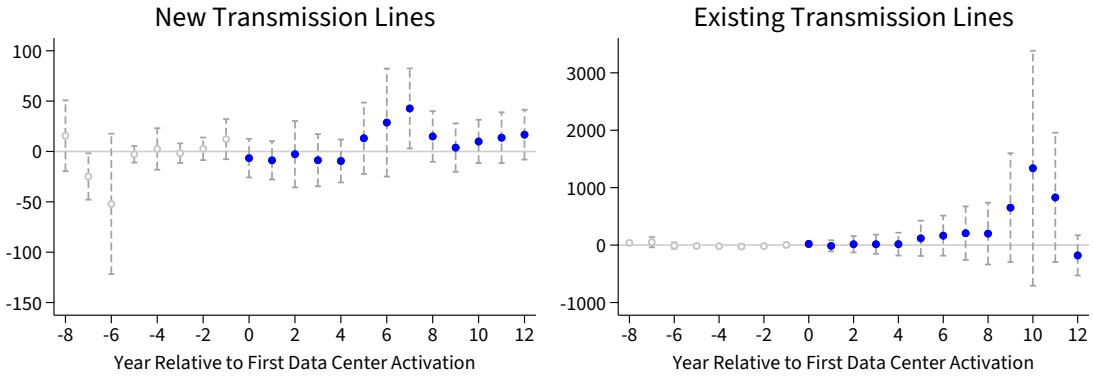
(f) Conductor Additions

Figure 4: Effect of Data Center Entry on Infrastructure Investment

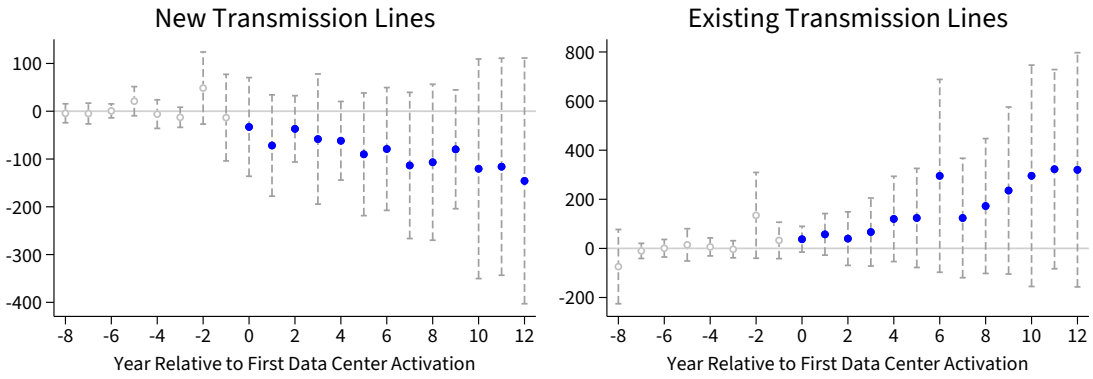
Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities (IOUs). Figures plot event-study coefficients and the 95% confidence intervals, capturing the difference in infrastructure investments for utilities with and without data centers before and after initial data center entry. The outcomes are (a) electric construction work in progress, (b) generator additions, (c) distribution station equipment additions, (d) transmission plant additions, (e) transmission operation & maintenance (O&M) expense, and (f) overhead conductor additions. All variables are in millions of 2024 U.S. dollars. Standard errors are clustered at the utility level.



(a) All Utilities



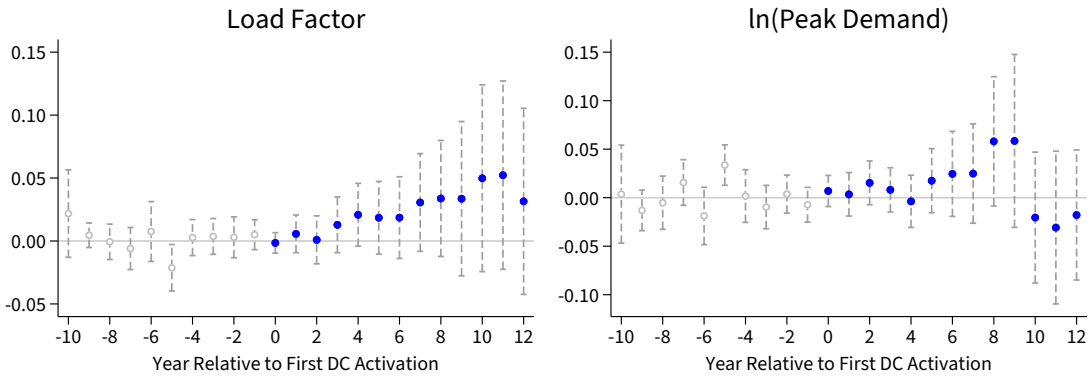
(b) Utilities Operating in Restructured States



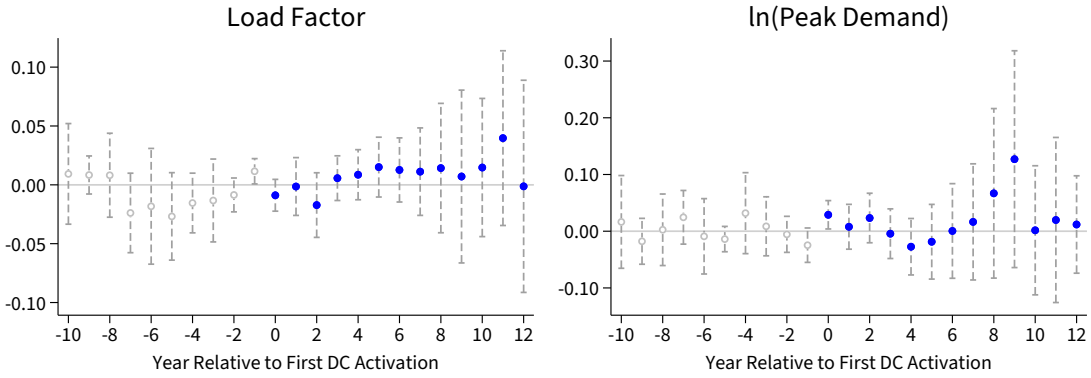
(c) Utilities Operating in Non-Restructured States

Figure 5: Effect of Data Center Entry on New and Existing Transmission Line Expenditures

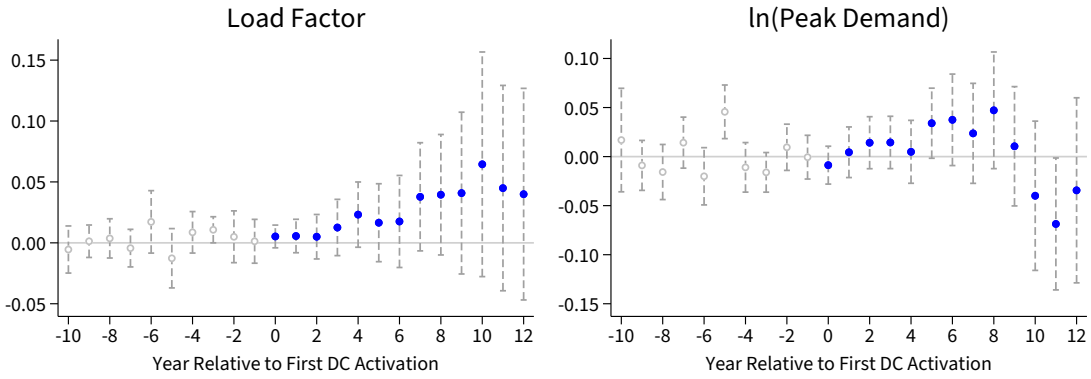
Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities (IOUs). Figures plot event-study coefficients and the 95% confidence intervals, capturing the difference in infrastructure investments for utilities with and without data centers before and after initial data center entry. The outcomes are expenditures on new transmission lines (left) and on existing transmission lines (right). Panel A reports estimates for all utilities. Panel B reports estimates for utilities operating only in restructured states, while Panel C reports estimates for utilities operating in non-restructured states. All variables are in millions of 2024 U.S. dollars. Standard errors are clustered at the utility level.



(a) All Utilities



(b) Utilities Operating in Restructured States



(c) Utilities Operating in Non-Restructured States

Figure 6: Effect of Data Center Entry on Load Factor and Peak Demand

Notes: The unit of observation is the utility-year. Figures plot event-study coefficients and the 95% confidence intervals. The left panels plot the system load factor, defined as a utility's total energy disposition divided by the product of peak demand and the number of hours in the year. The right panels plot the logarithm of system peak demand, measured in megawatts. Panel A reports estimates for all utilities. Panel B reports estimates for utilities operating only in restructured states, while Panel C reports estimates for utilities operating in non-restructured states. Standard errors are clustered at the utility level.

Main Text Tables

Table 1: Descriptive Statistics

	N	Mean	S.D.
A. Average Retail Electricity Prices ($/kWh)			
All	14,733	0.13	0.03
Residential	14,733	0.14	0.03
Commercial	14,733	0.13	0.03
Industrial	14,733	0.10	0.08
B. Data Centers			
PostDC	14,733	0.06	0.24
Cumulative capacity (MW)	14,733	4.97	44.51
Cumulative capacity (MW) conditional on entry	908	80.58	161.51
C. Infrastructure Investment ($ million)			
CWIP electric	1,590	327.48	621.38
Generator additions	1,590	33.16	148.36
Distribution station additions	1,590	22.57	33.54
Transmission plant additions	1,590	91.57	180.67
Transmission O&M	1,590	75.74	108.07
Conductor additions	1,590	15.89	38.63
Expenditure on new transmission lines	1,590	20.97	84.32
Expenditure on existing transmission lines	1,590	634.23	1029.92
D. Rate Cases			
Any case filed (0/1)	2,070	0.27	0.44
Any general case filed (0/1)	2,070	0.23	0.42
Any limited-issue rider case filed (0/1)	2,070	0.05	0.21
Number of cases filed	2,070	0.33	0.68
Mean rate case duration (months)	555	8.13	4.23
Requested rate change amount (USD billion)	2,070	0.03	0.08
Authorized rate change amount (USD billion)	2,070	0.01	0.05
Requested ROE (%)	462	10.34	0.58
Authorized ROE (%)	399	9.66	0.47
Requested return on rate base (%)	472	7.52	0.79
Authorized return on rate base (%)	355	7.09	0.76
Authorized rate base (USD billion)	336	3.36	4.81
E. Operational Characteristics			
Load factor	14,357	0.56	0.49
Peak demand (MW)	14,357	440.27	1361.44

Notes: Panels A, B, D are at the utility-state-year level. Panels C and E are at the utility-year level. Panels C and D only include investor-owned utilities. Dollar variables are in 2024 dollars.

Table 2: The Effect of Data Center Entry on Average Retail Prices

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
PostDC	0.027***	0.021***	0.028***	0.042**
	(0.009)	(0.008)	(0.009)	(0.016)
YMean ($/kWh)	0.14	0.15	0.14	0.10
# of Utilities	973	973	973	973
N	14,733	14,733	14,733	14,733

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall aggregated average treatment effect on the treated (ATT). The dependent variable is the log average retail price (total revenue divided by total sales, $/kWh), for all customers in column (1) and by customer class in columns (2)–(4). YMean reports the pre-treatment mean outcomes among treated units. Standard errors in parentheses are clustered at the utility-state level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Heterogeneity in Data Center Effects on Retail Price by Utility Ownership

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Investor-Owned Utilities				
PostDC	0.056***	0.052***	0.046***	0.080**
	(0.016)	(0.014)	(0.015)	(0.035)
YMean ($/kWh)	0.15	0.17	0.14	0.10
# of Utilities	900	900	900	900
N	13,585	13,585	13,585	13,585
B. Publicly-Owned Utilities				
PostDC	0.024*	0.017	0.017	0.027
	(0.015)	(0.015)	(0.015)	(0.020)
YMean ($/kWh)	0.13	0.15	0.14	0.10
# of Utilities	911	911	911	911
N	13,607	13,607	13,607	13,607
C. Cooperatives				
PostDC	-0.001	-0.007	0.018	0.014
	(0.013)	(0.010)	(0.013)	(0.022)
YMean ($/kWh)	0.14	0.15	0.14	0.11
# of Utilities	912	912	912	912
N	13,687	13,687	13,687	13,687

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price, defined as total revenue divided by total sales ($/kWh), for all customers in column (1) and separately by customer class in columns (2)–(4). Panel A reports estimates for investor-owned utilities, Panel B for public utilities, and Panel C for cooperatives. In each subsample analysis, the treatment group consists of utilities of the corresponding ownership type that eventually serve data centers, while the control group includes all utilities that never serve data centers, irrespective of ownership type. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Heterogeneity in Data Center Effects on Retail Price by Electricity Industry Structure

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Restructured States				
PostDC	0.038**	0.040**	0.029*	0.047
	(0.017)	(0.016)	(0.016)	(0.032)
YMean ($/kWh)	0.15	0.17	0.15	0.11
# of Utilities	239	239	239	239
N	3,421	3,421	3,421	3,421
B. Non-Restructured States				
PostDC	0.015	0.007	0.024**	0.027
	(0.010)	(0.009)	(0.010)	(0.017)
YMean ($/kWh)	0.13	0.14	0.13	0.10
# of Utilities	755	755	755	755
N	11,312	11,312	11,312	11,312

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price, defined as total revenue divided by total sales ($/kWh), for all customers in column (1) and separately by customer class in columns (2)–(4). Panel A reports estimates for utilities operating in restructured states, while Panel B reports estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: The Effect of Data Center Entry on Rate Case Activity

	Any Case Filed?			# Cases	Duration
	Any	General	Rider		
	(1)	(2)	(3)	(4)	(5)
A. Full Sample					
PostDC	0.084	0.104	-0.014	0.158	0.776
	(0.085)	(0.074)	(0.055)	(0.155)	(0.689)
YMean	0.330	0.277	0.068	0.429	8.246
ATT/YMean (%)	25.45	37.55	-20.59	36.83	9.41
N	2,070	2,070	2,070	2,070	555
B. Restructured States					
PostDC	0.265**	0.173*	0.122*	0.517**	-0.981*
	(0.110)	(0.097)	(0.070)	(0.234)	(0.549)
YMean	0.309	0.268	0.052	0.443	7.900
ATT/YMean (%)	85.76	64.55	234.62	116.70	-12.42
N	780	780	780	780	229
C. Non-Restructured States					
PostDC	-0.060	0.055	-0.115	-0.156	1.358
	(0.117)	(0.111)	(0.080)	(0.165)	(0.932)
YMean	0.351	0.287	0.085	0.415	8.561
ATT/YMean (%)	-17.09	19.16	-135.29	-37.59	15.86
N	1,290	1,290	1,290	1,290	326

Notes: The unit of observation is the utility-state-year and the sample only includes investor-owned utilities. The table reports the estimated overall average treatment effect on the treated (ATT). The outcomes are indicators for filing any rate case (column 1), any general rate case (column 2), and any limited-issue rider case (column 3); the number of cases filed (column 4); and the mean case duration in months, defined only for utility-state-years with at least one filed case (column 5). Panel A reports estimates for the full sample. Panel B reports estimates for utilities operating in restructured states, while Panel C reports estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: The Effect of Data Center Entry on Rate Case Requests and Authorizations

	Revenue Requirement ($ billion)		ROE (%)		Return on Rate Base (%)		Authorized Rate Base ($ billion)
	Requested (1)	Authorized (2)	Requested (3)	Authorized (4)	Requested (5)	Authorized (6)	(7)
A. Full Sample							
PostDC	0.017	-0.003	0.178*	-0.064	0.120	0.022	1.719***
	(0.017)	(0.013)	(0.092)	(0.123)	(0.104)	(0.202)	(0.460)
YMean	0.049	0.025	10.673	9.806	7.802	7.239	5.752
ATT/YMean (%)	34.69	-12.00	1.67	-0.65	1.54	0.30	29.89
N	2,070	2,070	462	399	472	355	336
B. Restructured States							
PostDC	0.040***	0.012	0.098	-0.182	0.100	-0.034	1.360*
	(0.014)	(0.013)	(0.159)	(0.212)	(0.120)	(0.378)	(0.696)
YMean	0.057	0.026	10.694	9.683	7.534	6.828	8.853
ATT/YMean (%)	70.18	46.15	0.92	-1.88	1.33	-0.50	15.36
N	780	780	200	188	199	150	134
C. Non-Restructured States							
PostDC	-0.002	-0.016	0.014	0.012	0.031	0.289	1.618**
	(0.029)	(0.021)	(0.066)	(0.056)	(0.130)	(0.232)	(0.631)
YMean	0.041	0.025	10.654	9.930	8.051	7.579	3.296
ATT/YMean (%)	-4.88	-64.00	0.13	0.12	0.39	3.81	49.09
N	1,290	1,290	262	211	273	205	202

Notes: The unit of observation is the utility-state-year and the sample only includes investor-owned utilities. The table reports the estimated overall average treatment effect on the treated (ATT). The outcomes are the requested and authorized change in the annual revenue requirement ($ billion) in columns (1)–(2), the requested and authorized return on equity (ROE, %) in (3)–(4), the requested and authorized return on rate base (%) in (5)–(6), and the authorized rate base ($ billion) in (7). The revenue-requirement outcomes are defined for every utility-state-year (zero when no change is decided), while the ROE, return-on-rate-base, and rate-base outcomes are defined only when a corresponding value is requested or decided. Panel A reports estimates for the full sample. Panel B reports estimates for utilities operating in restructured states, while Panel C reports estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix

Who Pays for Growth?
Evidence from Data Centers and the Grid

Robyn Meeks Jacquelyn Pless Zhiyuan Qi Zhenxuan Wang

A Additional Background **SI-1**

B Additional Details on Data and Analyses **SI-3**

C Additional Figures **SI-11**

D Additional Tables **SI-23**

A Additional Background

A.1 Rate Cases

State regulatory commissions carry out multiple functions that determine electric retail incidence and utility cost recovery, including both setting rates and overseeing utilities' major financial decisions. Much of this work occurs through rate cases, which are the formal regulatory proceedings that serve as the principal vehicle for any significant change to prices. The process usually begins when a utility applies for a rate increase, arguing that its existing rates no longer give it a fair opportunity to earn its allowed return. Utilities often file every few years while others go longer between cases; some states, however, mandate a regular schedule for rate cases. Once filed, rate cases move through a fairly standard sequence of events, which takes time and eventually culminates in the commission specifying new rates.[31]

Rate cases function to determine utilities' revenue requirements (by setting the utility's rate base and allowed rate of return), apportion that requirement across the different classes of customers, and determine tariff structures that will generate the utilities' allowable revenue. We detail these components here.

1. Determining the revenue requirement. This step is the core of a general rate case and is built around a "test year"— either a recently completed historical year or a forecasted future year — so that costs, investments, and sales all reflect the same period. The commission applies a basic formula to calculate the revenue requirement as rate base × rate of return + operating expenses. The rate base is the utility's net long-lived investment used to serve its customers (e.g., the original cost of plant in service less accumulated depreciation, adjusted for working capital, deferred taxes, and other approved items). The rate base includes items such as power plants, poles, wires, transformers, as well as office buildings and related office goods such as vehicles and computers. To be included in the rate base an investment must be both "used and useful" and prudently incurred. The rate of return is a weighted cost of capital across the utility's funding sources—common equity, preferred equity, and debt—with the cost of common equity being the most commonly contested element (Lazar, 2016). Operating expenses (i.e., labor, fuel, purchased power, taxes, and depreciation) are recovered on a dollar-for-dollar basis.

2. Allocating the revenue requirement across customer classes. Once the total revenue requirement is set, the commission decides how much each class of customers should contribute based on the usage characteristics of each class.[32] Costs are assigned according to cost causation,

[31] These steps include: the utility submits its proposed tariffs along with detailed supporting evidence; interested entities can join as parties, including the utility, the commission's staff, the consumer advocate, and any intervening entities that the commission admits, such as industrial, low-income, or environmental groups; the commission holds evidentiary hearings and separate sessions for public comment; lastly, the commission deliberates and issues a final order specifying the new rates and when they take effect.

[32] Notably, there is no national uniformity in how classes are defined — most utilities separate residential customers, and many add commercial, industrial, and other classes.

with some costs allocated by the number of customers, some by peak demand, and some by total energy consumption.[33] Allocation of costs can be driven by many factors, including individual judgment, politics, and economic development goals, amongst others.

3. Designing the rate within each customer class. The final step translates each class's allocated share into the actual prices customers pay, where essentially every rate element is an allocated portion of the revenue requirement divided by a billing determinant such as kWh, kW, or the number of customers. Rate design is normally handled within the general rate case. Addressing a single rate in isolation has been criticized with the concern that changing one customer's rate can shift costs onto others (Lazar, 2016).

A.2 Cost Recovery, Regulatory Lag, and Riders

The actual cost recovery for capital investments generally happens once the asset is "used and useful" and has been found to be prudently incurred, at which point the investment enters the rate base and the utility earns its allowed rate of return on the investment.

Given how rates are reset only periodically through rate cases, any investment a utility makes today does not begin generating revenue from customers until a later proceeding includes it. The gap in time between when a utility incurs costs and when those costs are reflected in the rates customers pay is referred to as "regulatory lag." Regulatory lag becomes important during infrastructure construction: an investment, particularly multi-year projects, could be on the utility's books for years before any rate case can incorporate the investment into rates. Mechanisms have been created to mitigate this lag.[34]

There is a faster, narrower alternative to waiting for a full rate case, and these mechanisms exist largely as a response to regulatory lag. For certain categories of spending—such as fuel and purchased power costs (through fuel adjustment clauses), and capital investments in infrastructure, smart-grids, or storm hardening (through infrastructure "trackers" or riders)—some utilities can recover costs between rate cases through single-issue filings that bypass the lag for those specific line items. Consumer advocates have been wary of these faster mechanisms, as they tend to address rising costs in isolation, whereas declining costs are only revisited in a full rate case (Lazar, 2016).

[33] There are two broad methodologies used to assign costs, both of which tend to assign higher per-kWh costs to residential and small business customers (who need more distribution investment and have "peakier" usage) than to industrial customers. These methodologies are: embedded cost studies, used by about 30 states, which divide the same historic (or future test-year) accounting costs that make up the revenue requirement among the classes so that the class totals sum to that requirement; and marginal cost studies, used by about 20 states, which apportion costs in proportion to the incremental cost of adding service.

[34] One mechanism is an allowance for funds used during construction (AFUDC), which accrues the cost of equity and debt tied up in the project during the build and adds it to the rate base when the asset enters service. This allows the utility to eventually earn on those carrying costs, but only later. Another is construction work in progress (CWIP), which places unfinished investment into rate base during construction so the utility earns a current return before the plant is finished. This effectively eliminates more of the lag.

B Additional Details on Data and Analyses

This appendix section provides additional details on the data sources and variable construction used in the main analysis as described in Section 4.1 and robustness checks.

B.1 Data for the Main Analyses

Data Centers. Our primary data source for data center entry is Aterio's *US Data Center Locations and Power Demand* product, a facility-level inventory of every active and proposed data center in the United States. As mentioned in the main paper, Aterio assembles these records from press releases and company websites, high-resolution satellite imagery, utility and interconnection filings, and permitting records from local and federal agencies. Importantly, every record is manually validated by analysts. Analysts assign each building a construction-completion percentage based on visual benchmarks in satellite imagery—from site preparation (5–15%) to equipment installation and commissioning (70–95%)—and log each observation with a timestamped description and analyst initials. Generator counts and equipment types are cross-checked against permitting documents and manufacturer specifications. Aterio additionally reconciles aggregate figures (total capacity, facility counts by state, and operator footprints) against 10-Q filings, utility disclosures, and permitting records. For facilities without power capacity reported by operators, capacity is estimated by an ensemble machine-learning model that combines facility size, generator counts identified from satellite imagery, operator, location, and construction stage. The model achieves an out-of-sample R^2 of 0.94 for hyperscalers and 0.97 for non-hyperscalers, and reports an 80% confidence interval for each estimate.

Retail Electricity Prices. We use data from EIA-861 to construct average retail electricity prices. As described in the paper, we use the bundled-service price. We cannot construct an all-in retail price for delivery-only customers from EIA-861 because the energy portion of their bill is paid to a competitive retail supplier, and EIA reports supplier sales only at the state-by-customer-class level rather than by the distribution utility serving each customer. The bundled-service price we use corresponds to the regulated default-service or standard-offer rate, which is set by the state public utility commission. In heavily restructured states, this price reflects a selected and shrinking subset of customers—those who never shopped, were returned to default service, or are explicitly on a regulated standard offer—rather than the market price faced by the typical customer.

Customer Classifications. EIA-861 reports retail sales and revenues by customer sector, including residential, commercial, industrial, and transportation customers. These sectors are based on utilities' reported customer classifications, which generally follow rate schedules and tariff definitions rather than a uniform end-use classification. This matters for data centers because facilities may be assigned to different reported sectors depending on the serving utility's tariff structure, voltage

level, demand size, and use of special contracts or large-load rate schedules. In practice, data centers may be reported as commercial, industrial, special contract, or large-load customers.

Starting in the year 2022, the EIA-861 instructions for Schedule 4 list "crypto mining and data centers" within the commercial sector. At the same time, EIA notes that customer classification follows the rate class rather than building activity, and a utility "may designate a large energy-intensive office or data center as an industrial customer."[35] We therefore interpret residential price effects as the cleanest evidence of pass-through to incumbent households. Commercial and industrial average price effects may combine pass-through to incumbent firms with the rates paid by data centers themselves.

Wholesale Electricity Prices. Our wholesale price data cover CAISO, ERCOT, ISO-NE, MISO, NYISO, PJM, and SPP. We use monthly averages of hourly day-ahead locational marginal prices and their energy, congestion, and loss components. The sample spans 2015–2024, except for CAISO, where coverage begins in 2016. Within each market, we only use pricing points with nonmissing observations for total LMP and all three components in every month of that market's sample period. This restriction holds the set of underlying pricing points constant over time and limits changes in measured prices caused by the entry or disappearance of pricing nodes.

We identify each utility-state pair's wholesale market affiliation using balancing-authority information from EIA-861. Where a utility can be matched to a corresponding price zone, we assign that zone's price. This approach applies to utility-associated transmission zones in PJM and default load aggregation points in CAISO. We verify that each zone belongs to the utility-state pair's market in the relevant year. A zonal price can therefore be shared by multiple utility-state pairs.

For utility-state pairs without a direct zone match, we construct prices from nearby physical pricing nodes. We first average prices across nodes at the same geographic coordinates so that locations with multiple node identifiers do not receive additional weight. For each county and market, we then select up to three of the nearest pricing locations within 300 kilometers of the county's Census internal point. We calculate a county-level price using weights proportional to the inverse of geographic distance. The selected locations and distance weights remain fixed over time within each county-market pair.

We aggregate county prices to the utility-state level within each organized market using 2010 county population shares within the utility's service territory. Service territories are defined by the counties reported in EIA-861 in 2010. Population serves as a proxy for the geographic distribution of electricity demand. For utility-state pairs operating in multiple organized markets, we combine market-specific prices using fixed sales shares calculated from EIA-861 over 2010–2012, normalized across the markets served in each year. When a baseline share is unavailable, we use the first

[35] See *Form EIA-861 Instructions*, U.S. Energy Information Administration, https://www.eia.gov/survey/form/eia_861/instructions.pdf; and U.S. Energy Information Administration, *Commercial Buildings Energy Consumption Survey (CBECS) Frequently Asked Questions*, https://www.eia.gov/consumption/commercial/faq.php.

observed share for that utility-state and market.

Annual wholesale prices are simple averages of the monthly measures, with at least ten months required for an annual observation. We apply the same spatial aggregation weights to total LMP and each component and deflate all four series using the annual CPI deflator used for retail prices, normalized to 2024. Wholesale prices are expressed in dollars per MWh. These measures capture wholesale price conditions associated with a utility's location and market participation. They do not measure its realized average procurement cost, which also depends on its purchasing arrangements and other charges.

B.2 Data for Robustness Checks on Confounding Factors

Matching to Grid Regions. We assign each utility-state to a grid region as follows. For utilities in organized wholesale markets, the region is the relevant ISO/RTO: CAISO, ERCOT, ISO-NE, MISO, NYISO, PJM, or SPP. Utilities outside organized wholesale markets are assigned to their NERC reliability region: FRCC, SERC, or WECC. Utility-specific trends are specified as linear trends. Grid-region-specific trends are specified as third-order polynomial trends. We do not use higher-order polynomial trends at the utility level because they would absorb too much of the identifying variation.

Renewable Generation and Storage Capacity. We construct annual measures of renewable generation and storage capacity using the U.S. Energy Information Administration's Form EIA-860, Annual Electric Generator Report, for 2010–2024. We use nameplate capacity from the annual operable generator inventories, which include operating, standby, and temporarily out-of-service units. We assign generators to counties using plant location information and aggregate capacity across the counties that each utility-state reports serving in EIA-861 in the corresponding year. These measures capture capacity located in the utility's service counties, regardless of ownership. A county served by multiple utilities contributes its full capacity to each utility's measure. Renewable capacity is the combined nameplate capacity of solar and wind generators. Panel A of Table A11 uses the log of one plus this capacity, measured in MW. Storage capacity includes batteries and other non-pumped-hydro storage, such as flywheels. Panel B uses storage capacity in MW. County-years with no recorded capacity of the relevant type receive a value of zero.

Wildfire Exposure. We construct historical wildfire exposure using the Monitoring Trends in Burn Severity (MTBS) Fire Occurrence Dataset, produced by the U.S. Geological Survey and the U.S. Forest Service. We use events from 1984–2009, before the start of our analysis sample, and retain wildfires and naturally ignited fires managed for resource objectives, excluding prescribed burns.[36] Each event reports its mapped area in acres and the geographic centroid of its fire perimeter. We

[36] MTBS primarily maps large fires, with thresholds of approximately 1,000 acres in the western United States and 500 acres in the eastern United States.

assign the event's entire mapped area to the county containing that centroid, using 2020 Census county boundaries. For each utility-state pair, we sum the mapped acreage across its fixed 2010 service counties and divide the total over 1984–2009 by 26 to obtain mean annual acreage. We then take the log of one plus this measure. Counties with no recorded qualifying events contribute zero. This variable measures historical wildfire exposure associated with the service counties.

Electric Vehicles. We obtain state-level battery-electric vehicle registrations from the U.S. Department of Energy's Alternative Fuels Data Center. The 2023 estimates are compiled from Experian registration data and cover light-duty all-electric vehicles, excluding plug-in hybrids. We assign each state's 2023 vehicle stock to all utility-state observations in that state. Panel D of Table A11 excludes observations in states whose stock exceeds 70,164 vehicles, the 75th percentile of this variable across utility-state-year observations in the analysis sample.

Electric Heating. We measure electric heating using the U.S. Census Bureau's American Community Survey (ACS) five-year estimates, Table B25040, House Heating Fuel, for the 2010–2024 vintages. Each vintage pools five years of survey responses. For each county, we divide the number of occupied housing units using electricity as their main heating fuel by the total number of occupied housing units. We aggregate these shares to the utility-state-year level using occupied housing units as weights across the utility's service counties. Equivalently, the utility-state measure is the total number of electrically heated occupied units in its service counties divided by their total occupied housing units.[37] Panel E of Table A11 uses this share in levels, with values between zero and one. The measure captures electricity as the main heating fuel and includes both heat pumps and electric resistance heating.

Smart Meter Deployment. We obtain annual advanced metering infrastructure (AMI) meter counts from EIA-861 for 2010–2024. For each utility-state-year, we divide the total number of AMI meters across customer classes by the number of distribution customers, defined as bundled-service plus delivery-only customers in the EIA-861 Sales to Ultimate Customers data. Energy-only customers are excluded from the denominator. Panel F of Table A11 uses the annual ratio in levels. Because customers may have multiple meters, this measure is AMI meters per distribution customer and is not capped at one.

B.3 Additional Details on Matched DID Design

The baseline estimates compare utility-state pairs that eventually serve data centers to those that never do. Although the Callaway and Sant'Anna (2021) estimator does not require treated and

[37] For Connecticut in 2022–2024, ACS county-level releases use planning regions. We reconstruct the eight historical counties used in our service-territory mapping by aggregating counts from the 169 towns using the Census Bureau's official geographic crosswalk. We use town counts from the corresponding ACS vintage in each year.

control units to be similar in levels, it does require parallel trends. Large differences in pre-period characteristics may therefore raise concerns that treated and never-treated utilities are structurally different in ways that could also be related to subsequent price dynamics. Panel A of Table A6 shows that such differences are present in the raw sample: before any data center entry, treated utilities served about 200,000 more customers, charged retail prices that were \$0.010/kWh higher, and operated in counties with roughly \$7,500 higher GDP per capita than never-treated utilities. All three differences are statistically significant at the 1% level.

We construct a more comparable control group using entropy balancing (Hainmueller, 2012). Entropy balancing assigns weights to never-treated utility-state pairs so that the weighted means of selected pre-period covariates among never-treated pairs exactly match the corresponding means among treated pairs, while keeping the weights as close as possible to uniform.[38] We balance on four pre-period characteristics measured over 2005–2007: average retail price, number of customers served, manufacturing establishments per capita, and GDP per capita. Because these characteristics are time-invariant pre-period measures, the balancing step is implemented once at the utility-state level. Treated utility-state pairs receive a weight of one. We then estimate the average treatment effect on the treated using the Callaway and Sant'Anna (2021) estimator on the weighted sample.

Panel B of Table A6 confirms that the reweighting achieves balance. After entropy balancing, none of the four covariates differs significantly between treated and never-treated utility-state pairs. The point estimates are also close to zero: the pre-period price gap falls from \$0.010/kWh to essentially zero, and the differences in customer count and GDP per capita also shrink dramatically.

Figure A4 plots the event-study coefficients estimated from the weighted sample. Table A7 reports the corresponding overall matched difference-in-differences estimates. If anything, the estimates are generally larger. This pattern suggests that observable pre-period imbalances likely attenuate the baseline estimates rather than inflate them.

B.4 Instrumental Variable Analysis

As a complementary approach, we also examine how cumulative data center capacity affects retail electricity prices, using interactions between historical local data-processing activity and aggregate investment by major technology firms as instrumental variables for cumulative capacity.

Data center location and expansion decisions may respond to local electricity prices, available infrastructure, and expected economic growth. These factors can also affect subsequent retail electricity prices. Our instruments combine historical differences in local data-processing activity with changes in aggregate annual capital expenditure by Google/Alphabet, Amazon, Meta, and Microsoft. The rationale is that locations with an established data-processing sector may be better positioned to attract additional capacity when these firms expand investment, due to existing

[38] Relative to nearest-neighbor or propensity-score matching, this approach discards no comparison units and achieves exact balance on the targeted moments, reducing researcher discretion over calipers and match ratios while preserving efficiency.

digital infrastructure, specialized workers and suppliers, and experience with facility development.

Historical Local Exposure. We measure historical data-processing activity using the U.S. Census Bureau's County Business Patterns (CBP) data for 2005–2007. We focus on NAICS 518210 (i.e., Data Processing, Hosting, and Related Services) and construct county-level averages of establishments and employment over these three years. County population comes from the Census Bureau's population estimates.

We map these county measures to utility-state pairs using a service-area crosswalk. For each utility-state pair i, let $\mathcal{C}_i$ denote the set of counties it reported serving in EIA-861. Let n_c denote the number of utilities serving county c in this crosswalk. We construct two exposure measures:

$$\begin{aligned} S_i^E &= 100{,}000 \frac{\sum_{c \in \mathcal{C}_i} \overline{\text{Est}}_c}{\sum_{c \in \mathcal{C}_i} \overline{\text{Pop}}_c}, \\ S_i^L &= \ln\left(1 + \sum_{c \in \mathcal{C}_i} \frac{\overline{\text{Emp}}_c}{n_c}\right), \end{aligned} \tag{A1}$$

where the bars denote averages over 2005–2007. The first measure is data-processing establishments per 100,000 residents across the utility-state's service counties. The second is the log of one plus data-processing employment allocated to the utility-state. For this allocation, we divide each county's employment equally among the utilities serving that county. The establishment-density measure uses the county totals directly in both its numerator and denominator.

CBP reports establishment counts but suppresses some employment values. We retain reported employment and impute suppressed values using establishment counts by employment-size category and representative employment levels within those categories. Where CBP provides an employment interval, we constrain the imputed value to that interval. County-years without a record for this sector are assigned zero establishments and employment. The establishment-density measure does not depend on the employment imputation.

Aggregate Capital Expenditure. We construct annual capital expenditure for Google / Alphabet, Amazon, Meta, and Microsoft from their annual filings, obtained through the Securities and Exchange Commission's EDGAR data. We sum expenditure across the four firms and express the resulting series, K_t, in nominal billions of dollars. Fiscal-year expenditure is assigned to the calendar year in which the fiscal year ends.[39] This measure captures aggregate corporate capital expenditure worldwide, including investment outside data centers.

For each exposure measure $k \in \{E, L\}$, we interact historical local exposure with standardized

[39] Microsoft's fiscal year ends in June. We use Google Inc.'s series through 2012 and Alphabet's historical series from 2013 onward. For Amazon, the earlier expenditure series is reported net of incentives, while the later series is reported gross. We align the earlier series by adding the 2016 overlap difference of $1.067 billion to observations through 2015.

annual capital expenditure:

$$W_{it}^k = S_i^k \frac{K_t - \overline{K}}{s_K}, \tag{A2}$$

where $\overline{K}$ and s_K are the mean and standard deviation of annual expenditure over 2010–2024. We then standardize each interaction over the estimation sample:

$$Z_{it}^k = \frac{W_{it}^k - \overline{W}^k}{s_{W^k}}. \tag{A3}$$

The same standardized instrument is used in both stages. These normalizations determine the units of the first-stage coefficients and do not change the two-stage least-squares estimates. We estimate a separate model for each exposure measure, with one excluded instrument in each model.

Capacity, Population, and Estimation Sample. The endogenous regressor is the log of one plus cumulative data center capacity, measured in MW. We construct capacity using active facilities in the Aterio inventory with usable activation years and capacity information. Facility capacity is assigned to the utilities serving its county. We impose no 5 MW threshold in this construction.

We control for annual log population. Using the same fixed county crosswalk as the historical exposure measures, we construct

$$N_{it} = \sum_{c \in \mathcal{C}_i} \frac{\text{Pop}_{ct}}{n_c}, \tag{A4}$$

where Pop_{ct} is the Census Bureau's annual county population estimate. Thus, population is allocated equally across utilities serving the same county, and the regression control is $\ln N_{it}$.[40]

The estimation sample covers the contiguous United States from 2010 through 2024. We exclude EIA-861 short-form filers and require historical service-county information, the population control, and valid prices for all four customer categories. After removing fixed-effect singletons, the common sample contains 16,874 utility-state-year observations for 1,186 utility-state pairs. Unlike the entry-based analysis, this sample retains utility-states that already hosted data centers at the beginning of the sample period.

Empirical Specification and Identification. Let C_{it} denote cumulative data center capacity and define $d_{it} = \ln(1 + C_{it})$. For each instrument k, the first-stage equation is

$$d_{it} = \pi_k Z_{it}^k + \gamma_k \ln N_{it} + \alpha_i + \lambda_t + u_{it}, \tag{A5}$$

where α_i and λ_t are utility-state and year fixed effects. For customer category j, the price equation is

$$\ln P_{it}^j = \beta_{kj} d_{it} + \delta_{kj} \ln N_{it} + \mu_i + \tau_t + \varepsilon_{it}^j, \tag{A6}$$

[40] For Connecticut in 2020–2024, we reconstruct population for the eight historical counties used in the service-area crosswalk by aggregating official town-level population estimates.

where P_{it}^{j} is average retail electricity price, measured overall or for residential, commercial, or industrial customers. We estimate equation (A6) by two-stage least squares, using Z_{it}^{k} as the excluded instrument for d_{it}. Both stages include the same population control and fixed effects. Standard errors are clustered at the utility-state level.

Utility-state fixed effects absorb permanent differences across service territories, while year fixed effects absorb common annual changes. Identification comes from differential changes in capacity across utility-states with different historical exposure as aggregate capital expenditure varies over time. The exclusion restriction requires that, conditional on the fixed effects and population control, this interaction affects retail prices only through data center capacity. In particular, historical data-processing activity must not proxy for other changes in local electricity demand or supply costs that co-vary with aggregate capital expenditure.

Results Appendix Table A8 reports the first-stage results. Both instruments positively predict cumulative data center capacity, with large F-stats suggesting strong predictive power.Appendix Table A9 reports the 2SLS estimates. Using the establishment-density instrument, the coefficient on log of one plus cumulative capacity is 0.073 for overall retail prices, 0.056 for residential prices, 0.043 for commercial prices, and 0.139 for industrial prices. Using the log-employment instrument, the corresponding estimates are 0.078, 0.071, 0.062, and 0.108. All estimates are positive and statistically significant at the 5% level.

Because capacity enters as $\ln(1 + C_{it})$, the implied elasticity of price with respect to capacity at a positive capacity level C is $\beta_{kj} C/(1 + C)$. For capacity levels large relative to one MW, the overall price estimates therefore approximate elasticities of 0.073 and 0.078.

C Additional Figures

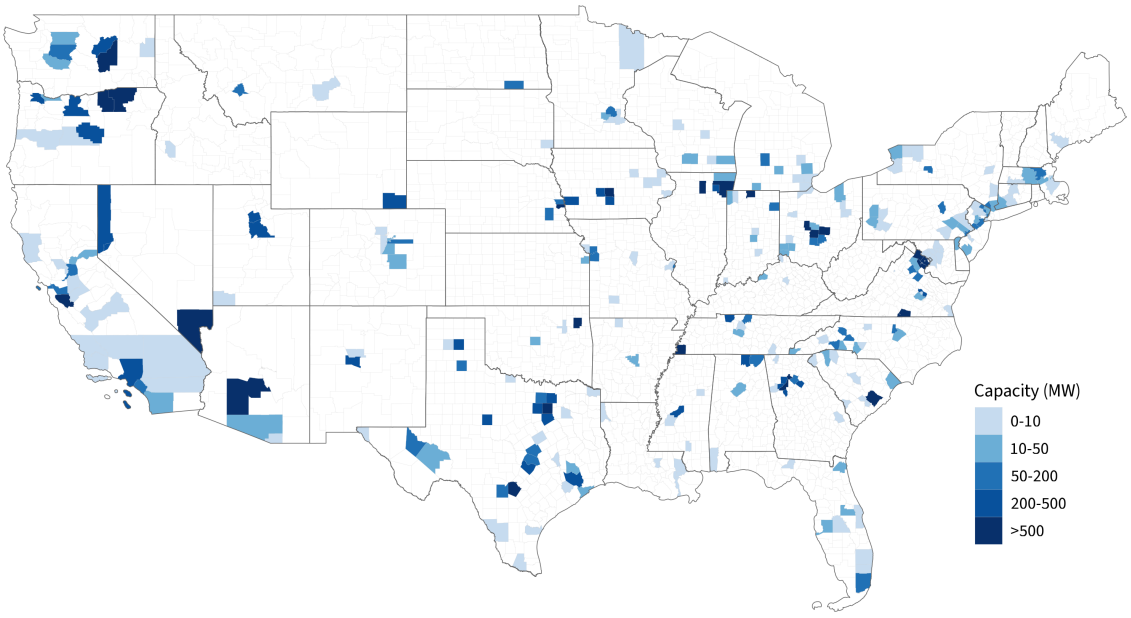
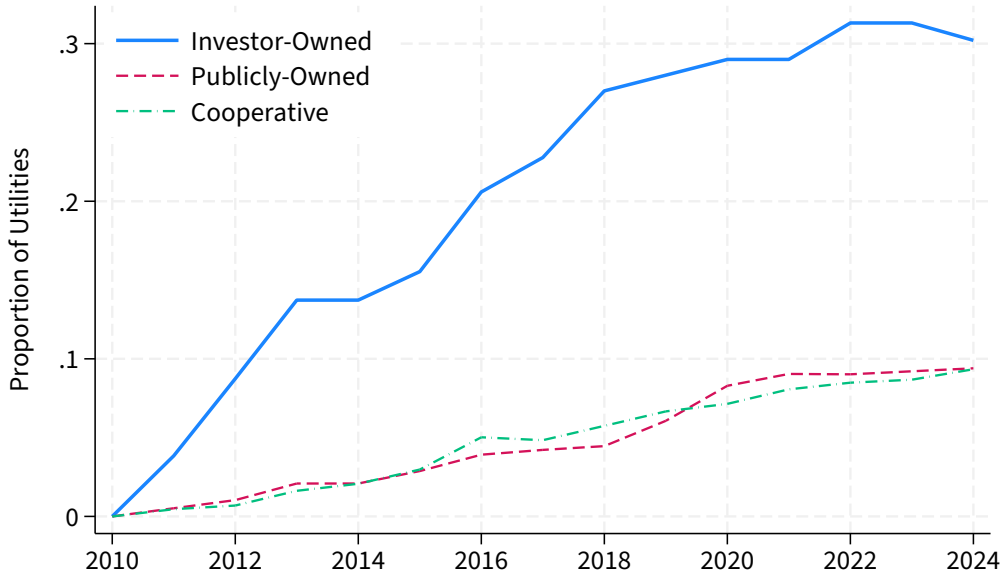
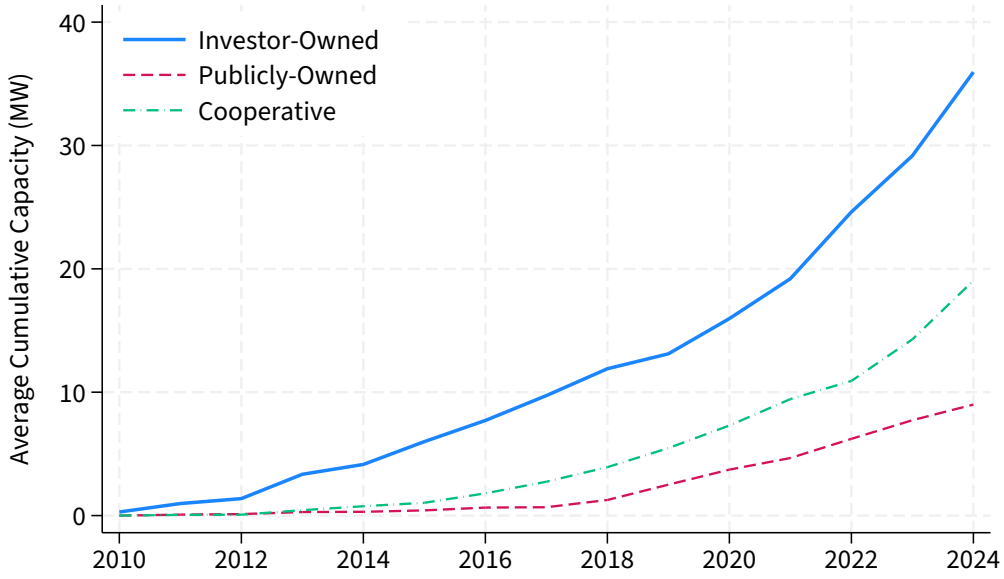


Figure A1: Data Center Capacity by County for Facilities Activated by the End of 2025

Notes: This figure shows total data center capacity (MW) by county in the contiguous United States. It includes all facilities that are activated by the end of 2025.



(a) Proportion of Utilities with Data Centers Over Time



(b) Average Cumulative Data Center Capacity by Utility Ownership

Figure A2: Data Center Entry and Expansion by Utility Ownership

Notes: Panel (a) plots the share of utility-states in the sample with at least one active data center in their service territory in each year. Panel (b) plots average cumulative active data center capacity (MW) per utility-state in the sample. Both outcomes are reported separately for investor-owned utilities, publicly owned utilities, and cooperatives.

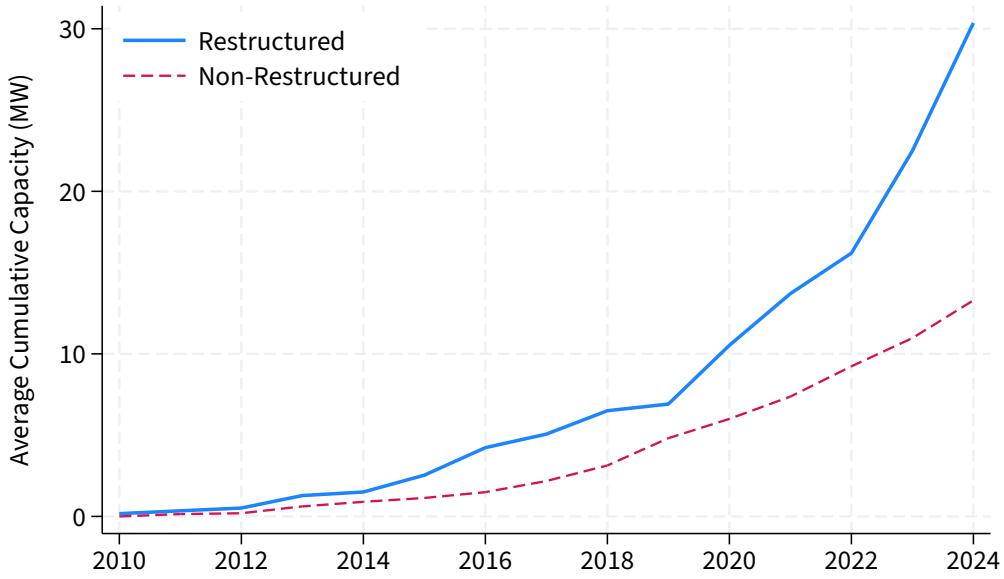


Figure A3: Average Cumulative Data Center Capacity by Electricity Industry Structure

Notes: The figure plots average cumulative active data center capacity (MW) per utility-state by year in our sample, separately for restructured and non-restructured states.

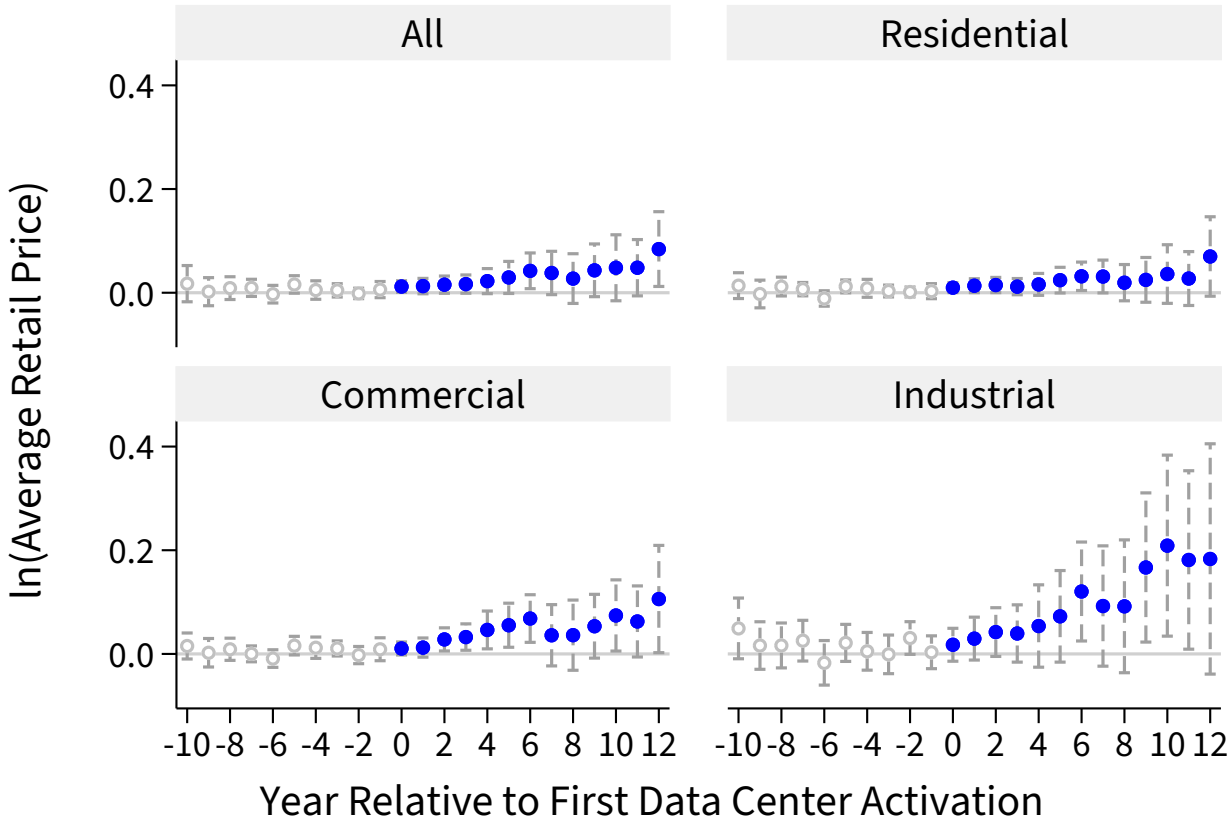
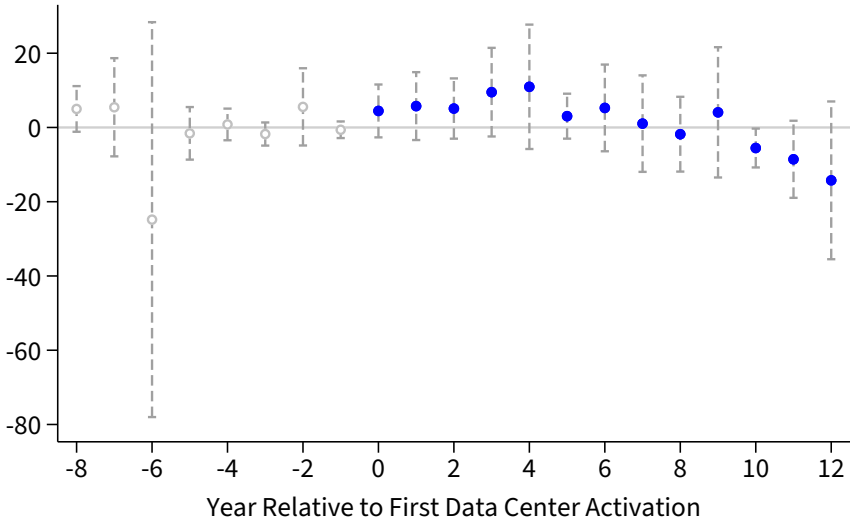
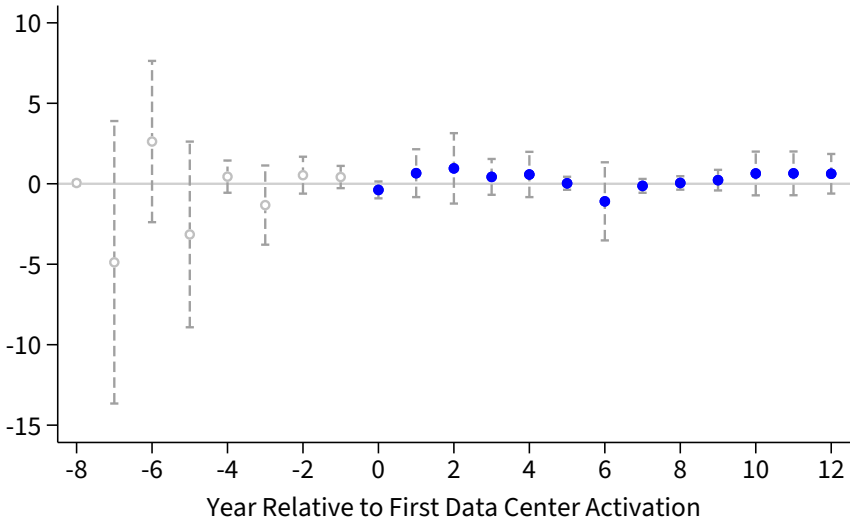


Figure A4: Dynamic Impact of Data Center Entry on Average Retail Prices – Matched DID Design via Entropy Balancing

Notes: We first reweight never-treated utility-state pairs by entropy balancing (Hainmueller, 2012) so that their pre-period means match those of the treated pairs, then estimate the overall average treatment effect on the treated (ATT) using the Callaway and Sant'Anna (2021) estimator with these weights. Balancing is performed once on four pre-treatment (2005–2007) characteristics measured before any data center entry: the average retail price, the number of customers, manufacturing establishments per capita, and GDP per capita. Treated units receive a weight of one. The unit of observation is the utility-state-year. Figures plot event-study coefficients and 95% confidence intervals. The first panel shows estimates for average retail rates across all customer classes, whereas the following three panels examine average rates for residential, commercial, and industrial customers separately. Standard errors are clustered at the utility-state level.



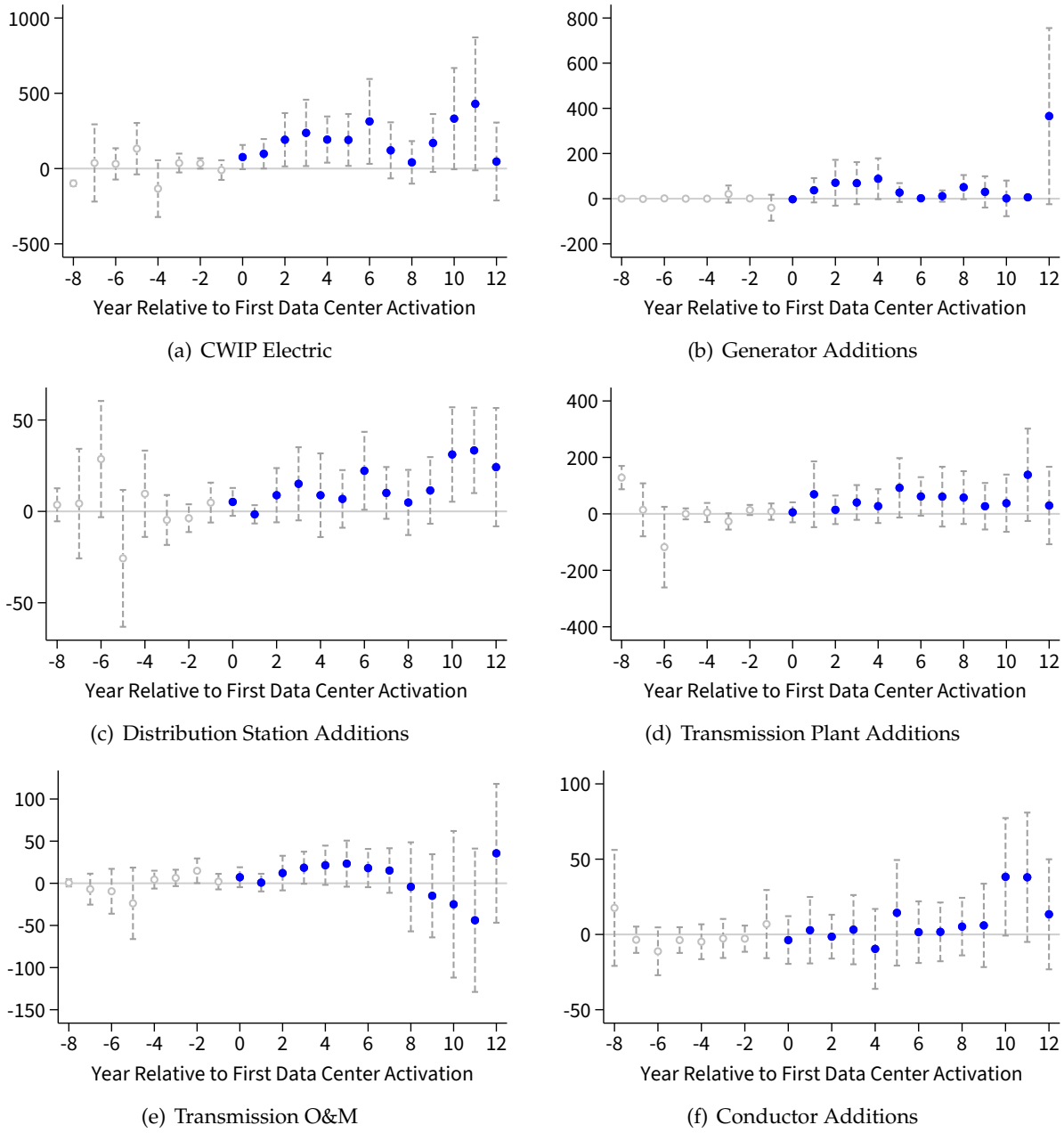
(a) CWIP Common



(b) CWIP Other

Figure A5: Placebo Tests on Construction Work in Progress

Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities (IOUs). Figures plot event-study coefficients and the 95% confidence intervals. Panel (a) plots common construction work in progress, i.e., projects that support multiple utility functions simultaneously such as IT systems, administrative offices, shared service centers, general office equipment, etc. Panel (b) plots other construction work in progress, i.e. projects attributable to other utility operations. All variables are in millions of 2024 U.S. dollars. Standard errors are clustered at the utility level.



(a) CWIP Electric

(b) Generator Additions

(c) Distribution Station Additions

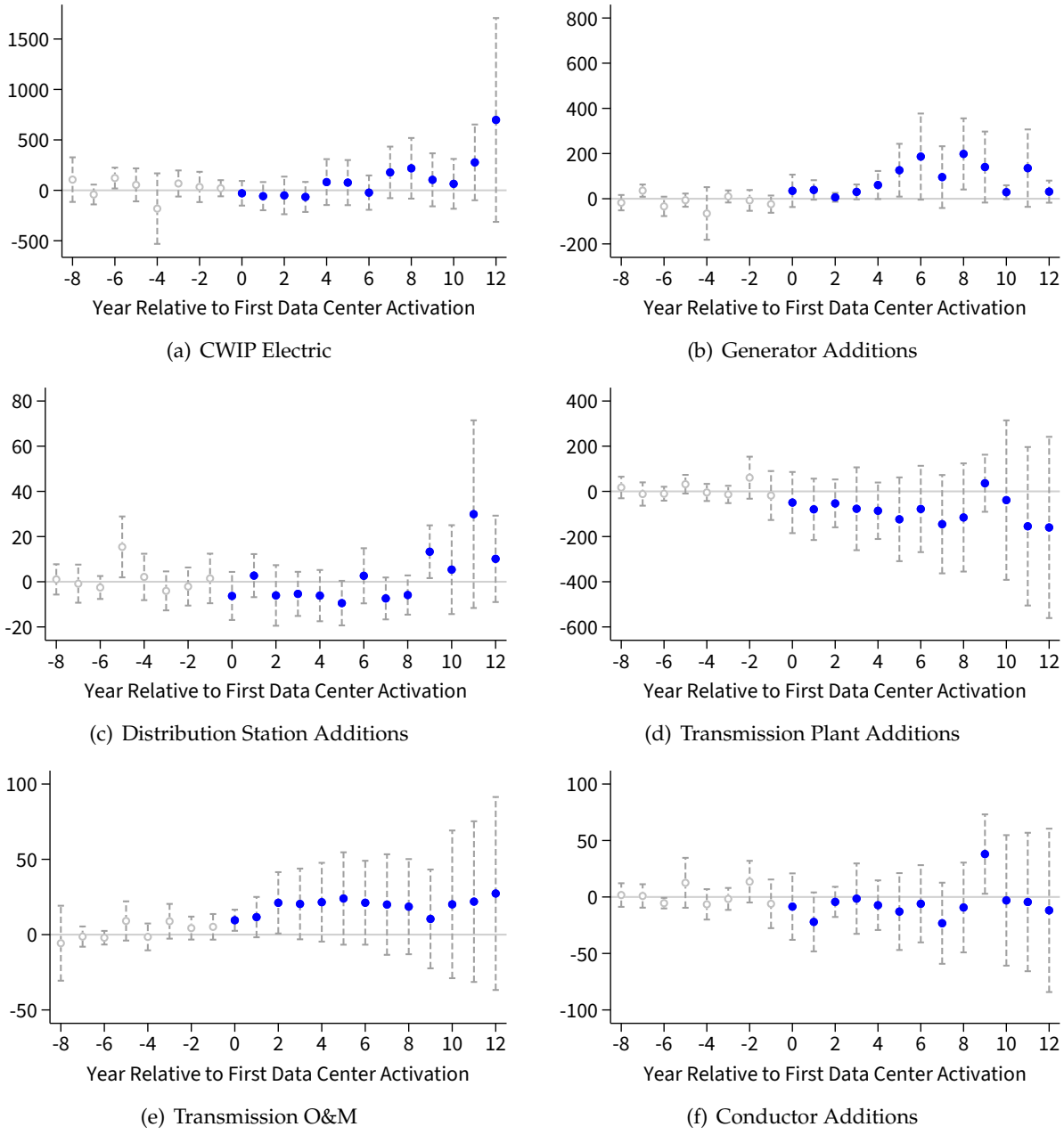
(d) Transmission Plant Additions

(e) Transmission O&M

(f) Conductor Additions

Figure A6: Effect of Data Center Entry on Infrastructure Investment for Utilities in Restructured Markets

Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities (IOUs). Figures plot event-study coefficients and the 95% confidence intervals using the sample of utilities that only operate in restructured states. The outcomes are (a) electric construction work in progress, (b) generator additions, (c) distribution station equipment additions, (d) transmission plant additions, (e) transmission operation & maintenance (O&M) expense, and (f) overhead conductor additions. All variables are in millions of 2024 U.S. dollars. Standard errors are clustered at the utility level.



(a) CWIP Electric

(b) Generator Additions

(c) Distribution Station Additions

(d) Transmission Plant Additions

(e) Transmission O&M

(f) Conductor Additions

Figure A7: Effect of Data Center Entry on Infrastructure Investment for Utilities in Non-Restructured Markets

Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities (IOUs). Figures plot event-study coefficients and the 95% confidence intervals using the sample of utilities that operate in at least one non-restructured state. The outcomes are (a) electric construction work in progress, (b) generator additions, (c) distribution station equipment additions, (d) transmission plant additions, (e) transmission operation & maintenance (O&M) expense, and (f) overhead conductor additions. All variables are in millions of 2024 U.S. dollars. Standard errors are clustered at the utility level.

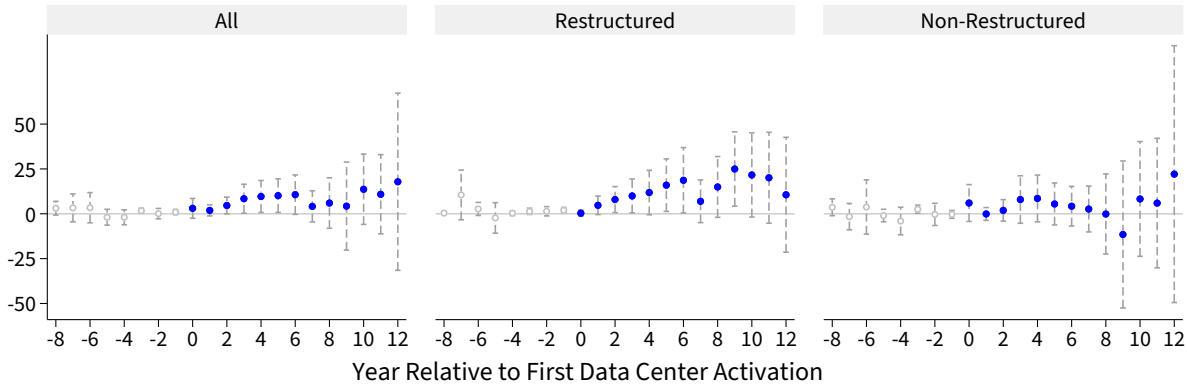


Figure A8: Effect of Data Center Entry on Utilities' Customer Advances for Construction

Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities (IOUs). Figures plot event-study coefficients and the 95% confidence intervals. The left panel plots estimates for all utilities. The middle panel plots estimates for utilities operating only in restructured states, while the right panel plots estimates for utilities operating in non-restructured states. The outcome is customer advances for construction (CAC), i.e., the upfront cash contributions customers make toward infrastructure built to serve them. All variables are in millions of 2024 U.S. dollars. Standard errors are clustered at the utility level.

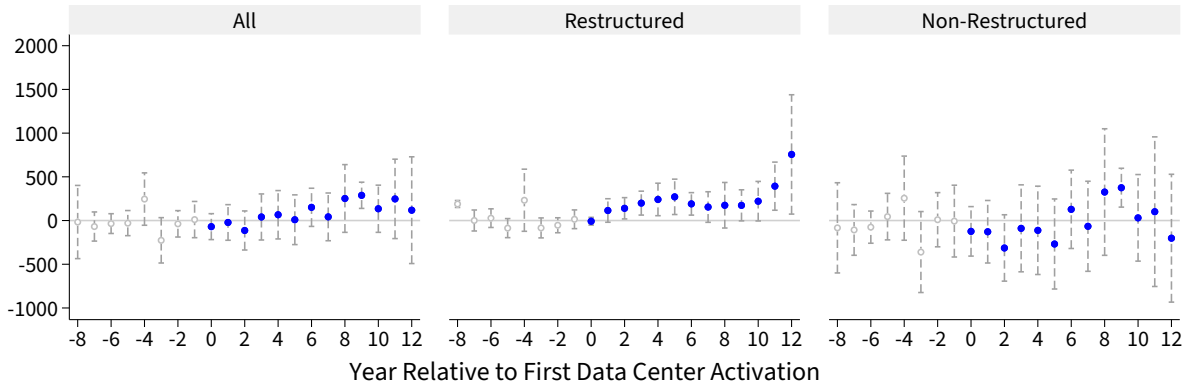


Figure A9: Effect of Data Center Entry on Additions to Electric Plants in Service

Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities (IOUs). Figures plot event-study coefficients and the 95% confidence intervals. The left panel plots estimates for all utilities. The middle panel plots estimates for utilities operating only in restructured states, while the right panel plots estimates for utilities operating in non-restructured states. The outcome is the addition to electric plant in service. All variables are in millions of 2024 U.S. dollars. Standard errors are clustered at the utility level.

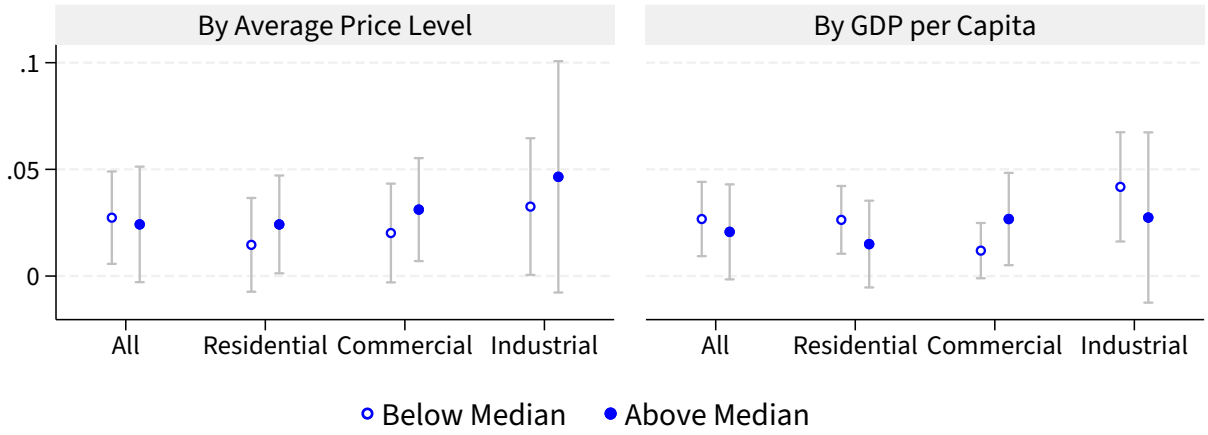


Figure A10: Distributional Consequences

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price for all customers and separately by customer class. Within each customer class, the sample is split at the median into "High" (above-median) and "Low" (at-or-below-median) subsamples. The left panel splits the sample by the utility-state's pre-treatment average retail price. The right panel splits the sample by pre-treatment county GDP per capita, averaged across counties in the utility-state's service territory. Both variables are measured using data from 2005–2007. Standard errors are clustered at the utility-state level. The corresponding estimates are reported in Table A4.

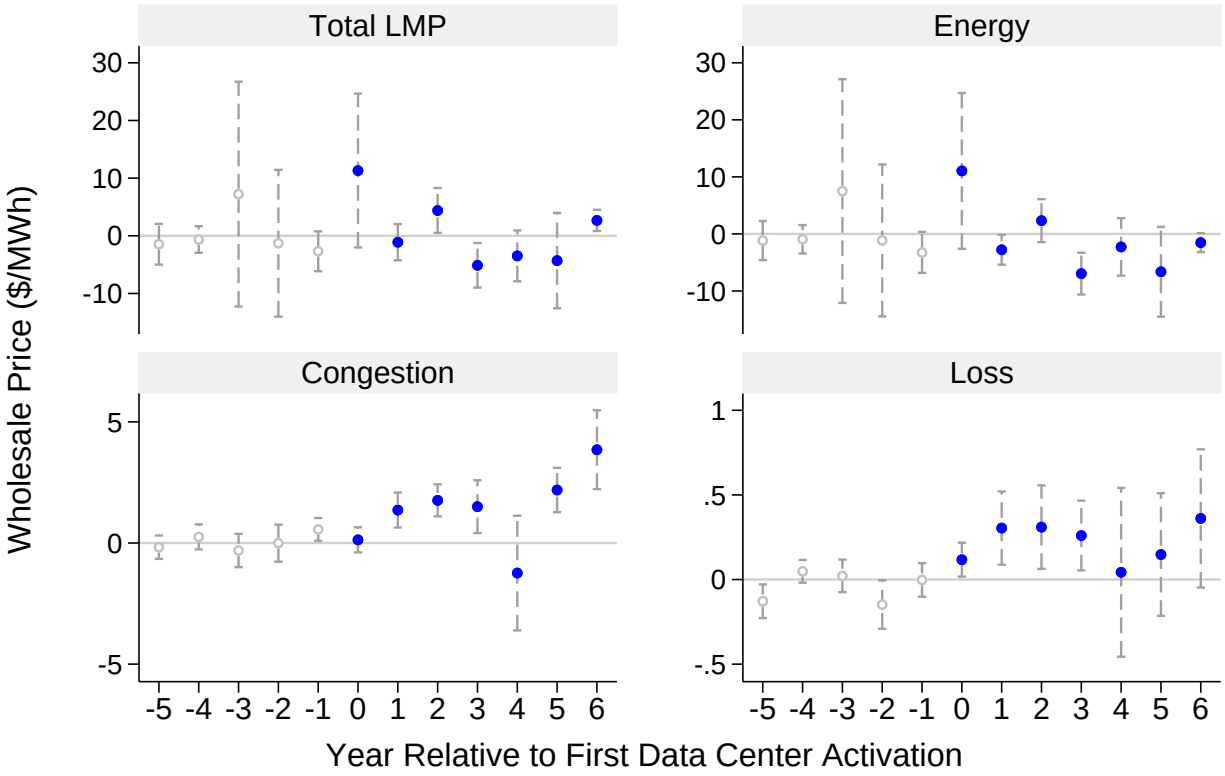


Figure A11: Dynamic Impact of Data Center Entry on Wholesale Prices

Notes: The unit of observation is the utility-state-year. The figure plots event-study coefficients and 95% confidence intervals for wholesale price exposures constructed from day-ahead nodal prices, using never-treated utility-state pairs as the comparison group. The first panel reports total locational marginal prices, whereas the following three panels examine the energy, congestion, and loss components separately. All four panels use the same sample. Outcomes are measured in levels and expressed in 2024 dollars per MWh. Standard errors are clustered at the utility-state level.

D Additional Tables

Table A1: Descriptive Statistics by Electricity Industry Structure

	Restructured			Non-Restructured		
	N	Mean	S.D.	N	Mean	S.D.
A. Average Retail Electricity Prices ($/kWh)						
All	3,421	0.14	0.04	11,312	0.12	0.02
Residential	3,421	0.16	0.05	11,312	0.14	0.02
Commercial	3,421	0.15	0.04	11,312	0.13	0.03
Industrial	3,421	0.11	0.05	11,312	0.10	0.09
B. Data Centers						
PostDC	3,421	0.10	0.30	11,312	0.05	0.22
Cumulative capacity (MW)	3,421	7.94	66.60	11,312	4.07	35.16
Cumulative capacity (MW) conditional on entry	351	77.36	194.83	557	82.61	136.55
C. Infrastructure Investment ($ million)						
CWIP electric	570	256.19	423.50	1,020	367.32	705.35
Generator additions	570	14.52	80.46	1,020	43.57	174.37
Distribution station additions	570	26.95	39.13	1,020	20.13	29.71
Transmission plant additions	570	86.32	139.16	1,020	94.50	200.15
Transmission O&M	570	83.57	124.42	1,020	71.37	97.53
Conductor additions	570	15.61	38.53	1,020	16.04	38.70
Expenditure on new transmission lines	570	12.21	45.31	1,020	25.87	99.37
Expenditure on existing transmission lines	570	544.47	932.57	1,020	684.39	1077.70
D. Rate Cases						
Any case filed (0/1)	780	0.29	0.46	1,290	0.25	0.44
Any general case filed (0/1)	780	0.25	0.43	1,290	0.22	0.41
Any limited-issue rider case filed (0/1)	780	0.05	0.23	1,290	0.04	0.21
Number of cases filed	780	0.42	0.88	1,290	0.28	0.51
Mean rate case duration (months)	229	8.25	4.39	326	8.04	4.12
Requested rate change amount ($ billion)	780	0.03	0.10	1,290	0.02	0.07
Authorized rate change amount ($ billion)	780	0.02	0.05	1,290	0.01	0.04
Requested ROE (%)	200	10.30	0.67	262	10.37	0.50
Authorized ROE (%)	188	9.55	0.51	211	9.75	0.41
Requested return on rate base (%)	199	7.45	0.79	273	7.57	0.79
Authorized return on rate base (%)	150	6.95	0.70	205	7.20	0.79
Authorized rate base ($ billion)	134	4.85	6.67	202	2.37	2.57
E. Operational Characteristics						
Load factor	3,159	0.52	0.16	11,198	0.57	0.55
Peak demand (MW)	3,159	569.42	1540.83	11,198	403.84	1304.16

Notes: Panels A, B, and D use utility-state-year observations. Panels C and E use utility-year observations. Panels C and D include IOUs only. Dollar values are in 2024 dollars.

Table A2: The Effect of Data Center Entry on Average Retail Prices by Subsequent Entry Status

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Without Subsequent Entry				
PostDC	0.017	0.010	0.008	0.039
	(0.012)	(0.011)	(0.010)	(0.026)
YMean ($/kWh)	0.14	0.16	0.14	0.11
# of Utilities	925	925	925	925
N	13,870	13,870	13,870	13,870
B. With Subsequent Entry				
PostDC	0.033***	0.029***	0.041***	0.044**
	(0.013)	(0.011)	(0.012)	(0.021)
YMean ($/kWh)	0.13	0.15	0.14	0.10
# of Utilities	924	924	924	924
N	13,936	13,936	13,936	13,936

Notes: The unit of observation is the utility-state-year. The table reports aggregated average treatment effects on the treated (ATT). Panel A includes treated utility-states with no additional data center entries after their initial entry year through 2024. Panel B includes those with subsequent entries during this period. Both panels retain the baseline never-treated control group. Outcomes are log average retail electricity prices. YMean reports the pre-treatment mean price among treated units, in 2024 dollars per kWh. Standard errors in parentheses are clustered at the utility-state level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Heterogeneity in Data Center Effects on Retail Price by Utility Ownership and Industry Structure

Dep. Var.:	All		Residential		Commercial		Industrial	
ln(Avg. Retail Price)	Restr.	Non-Restr.	Restr.	Non-Restr.	Restr.	Non-Restr.	Restr.	Non-Restr.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Investor-Owned Utilities								
PostDC	0.061**	0.044**	0.061***	0.040***	0.045*	0.042**	0.075	0.068**
	(0.025)	(0.018)	(0.023)	(0.014)	(0.025)	(0.016)	(0.068)	(0.026)
YMean ($/kWh)	0.17	0.12	0.19	0.15	0.16	0.13	0.12	0.09
# of Utilities	212	705	212	705	212	705	212	705
N	3,024	10,561	3,024	10,561	3,024	10,561	3,024	10,561
B. Publicly-Owned Utilities								
PostDC	0.025	0.019	0.035	0.003	0.020	0.012	0.031	0.014
	(0.038)	(0.012)	(0.036)	(0.012)	(0.038)	(0.011)	(0.041)	(0.022)
YMean ($/kWh)	0.14	0.13	0.16	0.14	0.15	0.13	0.11	0.10
# of Utilities	208	718	208	718	208	718	208	718
N	2,962	10,645	2,962	10,645	2,962	10,645	2,962	10,645
C. Cooperatives								
PostDC	0.019	-0.014	0.015	-0.019	0.016	0.019	0.025	0.003
	(0.019)	(0.017)	(0.017)	(0.013)	(0.015)	(0.019)	(0.017)	(0.033)
YMean ($/kWh)	0.13	0.14	0.16	0.15	0.15	0.14	0.10	0.11
# of Utilities	211	720	211	720	211	720	211	720
N	2,999	10,688	2,999	10,688	2,999	10,688	2,999	10,688

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price, defined as total revenue divided by total sales ($/kWh), for all customers and separately by customer class. Panel A reports estimates for investor-owned utilities, Panel B for public utilities, and Panel C for cooperatives. In each subsample analysis, the treatment group consists of utilities of the corresponding ownership type that eventually serve data centers, while the control group includes all utilities that never serve data centers, irrespective of ownership type. Odd-numbered columns report estimates for utilities operating in restructured states, while even-numbered columns report estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Distributional Consequences

Dep. Var.:	All		Residential		Commercial		Industrial	
ln(Avg. Retail Price)	High	Low	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: By Baseline Average Price Level								
PostDC	0.024*	0.027**	0.024**	0.015	0.031**	0.020*	0.046*	0.033**
	(0.014)	(0.011)	(0.012)	(0.011)	(0.012)	(0.012)	(0.028)	(0.016)
YMean ($/kWh)	0.15	0.12	0.17	0.14	0.15	0.13	0.11	0.09
# of Utilities	508	469	508	469	508	469	508	469
N	7,315	7,326	7,315	7,326	7,315	7,326	7,315	7,326
B: By GDP per Capita								
PostDC	0.021*	0.027***	0.015	0.026***	0.027**	0.012*	0.027	0.042***
	(0.011)	(0.009)	(0.010)	(0.008)	(0.011)	(0.007)	(0.020)	(0.013)
YMean ($/kWh)	0.14	0.13	0.16	0.15	0.14	0.14	0.11	0.10
# of Utilities	485	510	485	510	485	510	485	510
N	7,342	7,391	7,342	7,391	7,342	7,391	7,342	7,391

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price for all customers in columns (1)–(2), and separately by customer class in columns (3)–(8): residential in columns (3)–(4), commercial in columns (5)–(6), and industrial in columns (7)–(8). Within each pair of columns, the sample is split at the median into "High" (above-median) and "Low" (at-or-below-median) subsamples. Panel A splits the sample by the utility-state's pre-treatment average retail price. Panel B splits the sample by pre-treatment county GDP per capita, averaged across counties in the utility-state's service territory. Both variables are measured using data from 2005–2007. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Robustness Checks - Effect of Data Center Entry on Average Retail Prices

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Add utility-specific trend	0.026***	0.021***	0.028***	0.042**
	(0.009)	(0.008)	(0.009)	(0.016)
B. Add grid-region-specific trend	0.022**	0.019**	0.024***	0.029*
	(0.009)	(0.008)	(0.008)	(0.016)
C. Add pre average price × year FE	0.029***	0.023***	0.030***	0.046***
	(0.009)	(0.008)	(0.009)	(0.017)
D. Add pre number of customers × year FE	0.026***	0.016**	0.030***	0.056***
	(0.009)	(0.008)	(0.009)	(0.017)
E. Add pre total electricity sales × year FE	0.024***	0.013	0.028***	0.055***
	(0.009)	(0.008)	(0.009)	(0.017)
F. Add pre manufacturing establishments per capita × year FE	0.025***	0.020**	0.026***	0.039**
	(0.009)	(0.008)	(0.009)	(0.017)
G. Add pre population × year FE	0.037***	0.024***	0.042***	0.072***
	(0.009)	(0.009)	(0.009)	(0.017)
H. Add pre GDP per capita × year FE	0.031***	0.023***	0.032***	0.051***
	(0.009)	(0.008)	(0.009)	(0.017)
I. Drop earlier or later years of entry	0.025***	0.020**	0.027***	0.039**
	(0.009)	(0.008)	(0.009)	(0.018)
J. Drop utilities with partial cross-state exposure	0.028***	0.023***	0.029***	0.043***
	(0.009)	(0.008)	(0.009)	(0.016)
K. Exclude multi-state utilities	0.027***	0.021**	0.028***	0.042**
	(0.010)	(0.009)	(0.010)	(0.019)
L. Cluster the S.E. at the state level	0.027***	0.021***	0.028***	0.042***
	(0.009)	(0.008)	(0.008)	(0.014)
M. Cluster the S.E. at the utility level	0.027***	0.021***	0.028***	0.042**
	(0.009)	(0.008)	(0.009)	(0.016)

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated. The dependent variable is the log average retail electricity price, defined as total revenue divided by total sales ($/kWh). Panels A–B add utility-specific trends or grid-region-specific trends. Panels C–E add pre-period utility-state-level characteristics interacted with year fixed effects. Panels F–H add pre-period regional economic and demographic characteristics interacted with year fixed effects. Panel I drops treated utility-state pairs with first data center entry before 2012 or after 2022. Panels J–K exclude utilities that are exposed to data center entry only in a subset of their service states or operate in multiple states. Standard errors are clustered at the utility-state level, except in Panels L and M, where alternative clustering is indicated. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Covariate Balance Before and After Entropy Balancing

	Avg. Retail Price ($/kWh)	Number of Customers (10,000)	Manufacturing Establishments per Capita (# per 1,000 population)	GDP per Capita ($)
	(1)	(2)	(3)	(4)
A. Unweighted (Raw Sample)				
Ever Treated	0.010***	20.004***	-0.069	7494.563***
	(0.003)	(1.705)	(0.045)	(1367.940)
N	1,039	1,040	1,048	1,048
B. Entropy-Balanced (Weighted)				
Ever Treated	-0.000	-0.209	-0.005	98.503
	(0.002)	(2.868)	(0.023)	(1417.893)
N	1,039	1,040	1,041	1,041

Notes: This table assesses covariate balance between treated utility-state pairs (those that eventually serve data centers) and never-treated pairs, before and after entropy balancing (Hainmueller, 2012). The unit of observation is the utility-state pair, measured once using pre-period (2005–2007) values. Each cell reports the coefficient from a regression of the listed pre-period characteristic on an indicator for ever serving a data center. Panel A uses the raw (unweighted) sample. Panel B reweights never-treated pairs by their entropy-balancing weights, which are chosen so that the weighted never-treated means equal the treated means. The characteristics are the average retail price ($/kWh), the number of customers, manufacturing establishments per capita, and GDP per capita. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Robustness Checks - Matched DID via Entropy Balancing

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Full Sample	0.027**	0.021**	0.039**	0.074*
	(0.011)	(0.009)	(0.016)	(0.039)
N	14,661	14,661	14,661	14,661
B. Investor-Owned Utilities	0.061***	0.056***	0.063***	0.127**
	(0.019)	(0.015)	(0.024)	(0.061)
N	13,513	13,513	13,513	13,513
C. Publicly-Owned Utilities	0.020	0.013	0.023	0.051
	(0.017)	(0.016)	(0.020)	(0.040)
N	13,535	13,535	13,535	13,535
D. Cooperatives	-0.003	-0.009	0.028	0.037
	(0.015)	(0.012)	(0.019)	(0.038)
N	13,615	13,615	13,615	13,615
E. Restructured States	0.056**	0.051***	0.073**	0.183**
	(0.024)	(0.019)	(0.033)	(0.090)
N	3,386	3,386	3,386	3,386
F. Non-Restructured States	0.006	-0.001	0.018*	0.018
	(0.010)	(0.009)	(0.010)	(0.017)
N	11,275	11,275	11,275	11,275

Notes: This table reports estimates from a matched difference-in-differences design. We first reweight never-treated utility-state pairs by entropy balancing (Hainmueller, 2012) so that their pre-period means match those of the treated pairs, then estimate the overall average treatment effect on the treated (ATT) using the Callaway and Sant'Anna (2021) estimator with these weights. Balancing is performed once on four pre-treatment (2005–2007) characteristics measured before any data center entry: the average retail price, the number of customers, manufacturing establishments per capita, and GDP per capita. Treated units receive a weight of one. The unit of observation is the utility-state-year. The dependent variable is the log average retail electricity price, defined as total revenue divided by total sales ($/kWh), for all customers and separately by customer class. Panel A reports estimates using the full sample. Panel B reports estimates for investor-owned utilities, Panel C for public utilities, and Panel D for cooperatives. In each subsample analysis by ownership, the treatment group consists of utilities of the corresponding ownership type that eventually serve data centers, while the control group includes all utilities that never serve data centers, irrespective of ownership type. Panel E reports estimates for utilities operating in restructured states while Panel F reports estimates for utilities operating in non-restructured states. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: First-Stage Estimates for Data Center Capacity

Dep. Var.:	$\ln(1 + \text{Capacity})$	
	(1)	(2)
Data-processing establishments per 100k residents × CAPEX	0.146***	
	(0.021)	
$\ln(1 + \text{Data-processing employment}) \times \text{CAPEX}$		0.206***
		(0.018)
Observations	16,874	16,874
KP F-stats	48.38	124.80

Notes: First-stage estimates corresponding to the two-stage least-squares specifications in Table A9, using the 2010–2024 utility-state-year sample. The dependent variable is the log of one plus cumulative data center capacity in MW. Each column includes one instrument, constructed by interacting historical exposure to the data-processing and hosting sector in 2005–2007 with aggregate annual capital expenditure by Google/Alphabet, Amazon, Meta, and Microsoft. Exposure is establishments per 100,000 residents in Column (1) and the log of one plus employment allocated to the utility-state in Column (2). Each instrument is standardized to have mean zero and unit standard deviation within the estimation sample. Both specifications include log population and utility-state and year fixed effects. Standard errors in parentheses are clustered at the utility-state level. KP F denotes the Kleibergen–Paap rk Wald statistic from the corresponding two-stage least-squares specification. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: IV Estimates of Data Center Capacity Effects on Retail Prices

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Data-processing establishments per 100k residents $\times$ *CAPEX*				
$\ln(1 + \text{Capacity})$	0.073***	0.056***	0.043**	0.139***
	(0.022)	(0.019)	(0.021)	(0.039)
Observations	16,874	16,874	16,874	16,874
KP *F*-stats	48.38	48.38	48.38	48.38
B. $\log(1 + \textit{Data-processing employment}) \times \textit{CAPEX}$				
$\ln(1 + \text{Capacity})$	0.078***	0.071***	0.062***	0.108***
	(0.015)	(0.014)	(0.015)	(0.022)
Observations	16,874	16,874	16,874	16,874
KP *F*-stats	124.80	124.80	124.80	124.80

Notes: Two-stage least-squares estimates for 2010–2024. The unit of observation is the utility-state-year. Outcomes are log average retail electricity prices, overall and by customer class. The endogenous regressor is the log of one plus cumulative data center capacity in MW. Each panel uses one instrument: historical exposure to the data-processing and hosting sector in 2005–2007 interacted with aggregate annual capital expenditure by Google/Alphabet, Amazon, Meta, and Microsoft. Exposure is establishments per 100,000 residents in Panel A and the log of one plus employment allocated to the utility-state in Panel B. Each instrument is standardized to have mean zero and unit standard deviation within the estimation sample. All specifications include utility-state and year fixed effects and log population. Standard errors in parentheses are clustered at the utility-state level. KP *F* denotes the Kleibergen–Paap rk Wald statistic. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Robustness Checks - Omit Utility-States with Data Center Entry Before 2016

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Investor-Owned Utilities				
PostDC	0.048**	0.054***	0.052***	0.068*
	(0.021)	(0.013)	(0.019)	(0.038)
N	13,325	13,325	13,325	13,325
B. Publicly-Owned Utilities				
PostDC	0.004	0.004	0.001	-0.017
	(0.010)	(0.012)	(0.009)	(0.020)
N	13,442	13,442	13,442	13,442
C. Cooperatives				
PostDC	0.015	0.015	0.041**	-0.020
	(0.013)	(0.011)	(0.020)	(0.026)
N	13,488	13,488	13,488	13,488
D. Restructured States				
PostDC	0.039***	0.035***	0.023	0.053*
	(0.014)	(0.011)	(0.015)	(0.028)
N	3,161	3,161	3,161	3,161
E. Non-Restructured States				
PostDC	0.007	0.012	0.032**	-0.026
	(0.011)	(0.009)	(0.014)	(0.020)
N	10,948	10,948	10,948	10,948

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). We omit utility-state pairs with first data center entry before 2016 from the treated group. The dependent variable is the log average retail electricity price, defined as total revenue divided by total sales ($/kWh), for all customers and separately by customer class. Panel A reports estimates for investor-owned utilities, Panel B for public utilities, and Panel C for cooperatives. In each subsample analysis by ownership, the treatment group consists of utilities of the corresponding ownership type that eventually serve data centers, while the control group includes all utilities that never serve data centers, irrespective of ownership type. Panel D reports estimates for utilities operating in restructured states while Panel E reports estimates for utilities operating in non-restructured states. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Robustness Checks Addressing Alternative Explanations

Dep. Var.: ln(Avg. Retail Price)	All (1)	Residential (2)	Commercial (3)	Industrial (4)
A. Control for renewable energy capacity	0.027***	0.022***	0.028***	0.039**
	(0.009)	(0.008)	(0.009)	(0.016)
B. Control for storage capacity	0.027***	0.021***	0.028***	0.041**
	(0.009)	(0.008)	(0.009)	(0.017)
C. Control for wildfire exposure $\times$ linear time trend	0.027***	0.021***	0.028***	0.042**
	(0.009)	(0.008)	(0.009)	(0.017)
D. Exclude observations in states with high EV stocks	0.028***	0.020**	0.029***	0.042***
	(0.010)	(0.009)	(0.009)	(0.015)
E. Control for electric heating share	0.025***	0.022**	0.027***	0.032**
	(0.009)	(0.008)	(0.009)	(0.016)
F. Control for smart meters deployment	0.028***	0.022***	0.029***	0.043***
	(0.009)	(0.008)	(0.009)	(0.016)
G. Exclude California	0.027***	0.022***	0.028***	0.043**
	(0.009)	(0.008)	(0.009)	(0.017)
H. Exclude Texas and Virginia	0.025***	0.019**	0.026***	0.038**
	(0.009)	(0.009)	(0.009)	(0.017)

Notes: The unit of observation is the utility-state-year. The table reports overall average treatment effects on the treated. The dependent variable is the log average retail electricity price, defined as total revenue divided by total sales ($/kWh). Panel A controls for the log of one plus installed solar and wind capacity (MW); Panel B controls for storage capacity (MW), excluding pumped hydro. Both measures are summed over the utility's service counties. Panel C controls for pre-sample wildfire exposure interacted with a linear time trend. Wildfire exposure is the log of one plus mean annual burned acres during 1984–2009. Panel D excludes observations in states whose 2023 battery-electric vehicle stock exceeds the 75th percentile across analysis-panel observations. Panel E controls for the share of occupied housing units using electricity as their main heating fuel, calculated from ACS five-year estimates and weighted by occupied housing units across service counties. Panel F controls for annual advanced metering infrastructure (AMI) meters per distribution customer. Panels G–H exclude observations in California and in Texas and Virginia, respectively. Standard errors in parentheses are clustered at the utility-state level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: The Effect of Data Center Entry on Utilities' Infrastructure Investments

	CWIP Electric	Generator Additions	Distribution Station Additions	Transmission Plant Additions	Transmission O&M	Conductor Additions	New Transmission Lines	Existing Transmission Lines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. All Utilities								
PostDC	137.39**	63.57***	5.48	-17.30	12.65	-0.88	-33.84	163.39
	(59.31)	(20.69)	(3.44)	(45.32)	(8.95)	(7.30)	(30.56)	(107.83)
YMean	394.50	14.86	30.19	92.26	95.37	14.39	23.92	747.98
ATT/YMean (%)	34.83	427.79	18.15	-18.75	13.26	-6.12	-141.47	21.84
N	1,590	1,590	1,590	1,590	1,590	1,590	1,590	1,590
B. Utilities Operating in Restructured States								
PostDC	186.56***	41.95**	11.54**	48.61	8.59	5.13	6.45	189.48
	(69.32)	(20.84)	(5.50)	(31.37)	(11.02)	(8.14)	(10.30)	(190.67)
YMean	351.52	10.86	32.76	63.13	114.38	6.98	13.56	572.54
ATT/YMean (%)	53.07	386.28	35.23	77.00	7.51	73.50	47.57	33.09
N	570	570	570	570	570	570	570	570
C. Utilities Operating in Non-Restructured States								
PostDC	81.63	84.91**	0.52	-72.34	18.38	-3.55	-68.85	159.89
	(85.25)	(33.69)	(3.77)	(80.09)	(13.65)	(12.48)	(56.64)	(108.61)
YMean	434.20	18.55	27.82	119.17	77.81	21.23	33.48	910.07
ATT/YMean (%)	18.80	457.74	1.87	-60.70	23.62	-16.72	-205.65	17.57
N	1,020	1,020	1,020	1,020	1,020	1,020	1,020	1,020

Notes: The unit of observation is the utility-year, and the sample includes investor-owned utilities only. The table reports the estimated overall average treatment effect on the treated (ATT). Panel A uses all utilities. Panel B uses the sample of utilities that operate only in restructured states. Panel C uses the sample of utilities that operate in at least one non-restructured state. The outcomes are (1) electric construction work in progress, (2) generator additions, (3) distribution station equipment additions, (4) transmission plant additions, (5) transmission operation & maintenance (O&M) expense, (6) overhead conductor additions, (7) expenditures on new transmission lines, and (8) expenditures on existing transmission lines. All variables are in millions of 2024 U.S. dollars. YMean reports the pre-treatment mean outcomes among treated units. Standard errors clustered at the utility level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Heterogeneity in Data Center Effects on Retail Price by Grid Reserve Margin

Dep. Var.:	All		Residential		Commercial		Industrial	
ln(Avg. Retail Price)	High	Low	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Full Sample								
PostDC	0.019	0.036***	0.017	0.028**	0.022	0.035***	0.047	0.033*
	(0.014)	(0.012)	(0.012)	(0.011)	(0.013)	(0.011)	(0.029)	(0.018)
YMean ($/kWh)	0.14	0.13	0.15	0.15	0.15	0.14	0.11	0.10
# of Utilities	474	490	474	490	474	490	474	490
N	7,293	7,405	7,293	7,405	7,293	7,405	7,293	7,405
B. Restructured States								
PostDC	0.008	0.056***	0.007	0.059***	-0.012	0.057***	0.041	0.047*
	(0.028)	(0.020)	(0.026)	(0.019)	(0.026)	(0.020)	(0.060)	(0.028)
YMean ($/kWh)	0.17	0.14	0.19	0.16	0.18	0.14	0.13	0.10
# of Utilities	83	154	83	154	83	154	83	154
N	1,225	2,176	1,225	2,176	1,225	2,176	1,225	2,176
C. Non-Restructured States								
PostDC	0.005	0.023	0.006	0.007	0.026*	0.021*	0.017	0.022
	(0.012)	(0.015)	(0.011)	(0.013)	(0.015)	(0.012)	(0.026)	(0.023)
YMean ($/kWh)	0.13	0.13	0.14	0.15	0.14	0.13	0.10	0.10
# of Utilities	401	347	401	347	401	347	401	347
N	6,068	5,229	6,068	5,229	6,068	5,229	6,068	5,229

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price for all customers and separately by customer class. Within each pair of columns, the sample is split at the median into "High" (above-median) and "Low" (at-or-below-median) subsamples based on the utility-state's pre-sample balancing-authority reserve margin, measured over 2005–2007. Panel A reports estimates for the full sample. Panel B reports estimates for utilities operating in restructured states, while Panel C reports estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Heterogeneity in Data Center Effects on Retail Price by Grid Fossil Share

Dep. Var.:	All		Residential		Commercial		Industrial	
ln(Avg. Retail Price)	High	Low	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Full Sample								
PostDC	0.024*	0.028**	0.022*	0.021**	0.036**	0.025**	0.011	0.055**
	(0.014)	(0.011)	(0.013)	(0.010)	(0.016)	(0.010)	(0.022)	(0.022)
YMean ($/kWh)	0.13	0.14	0.15	0.16	0.14	0.14	0.10	0.11
# of Utilities	280	686	280	686	280	686	280	686
N	4,552	10,166	4,552	10,166	4,552	10,166	4,552	10,166
B. Restructured States								
PostDC	0.059*	0.022	0.063**	0.023	0.058**	0.010	0.031	0.049
	(0.031)	(0.021)	(0.029)	(0.019)	(0.028)	(0.020)	(0.038)	(0.045)
YMean ($/kWh)	0.14	0.15	0.17	0.17	0.15	0.15	0.10	0.12
# of Utilities	102	137	102	137	102	137	102	137
N	1,396	2,025	1,396	2,025	1,396	2,025	1,396	2,025
C. Non-Restructured States								
PostDC	0.017	0.011	0.017	-0.003	0.032*	0.017	-0.003	0.043*
	(0.015)	(0.013)	(0.014)	(0.010)	(0.018)	(0.011)	(0.023)	(0.024)
YMean ($/kWh)	0.13	0.13	0.15	0.14	0.13	0.14	0.10	0.10
# of Utilities	236	512	236	512	236	512	236	512
N	3,783	7,514	3,783	7,514	3,783	7,514	3,783	7,514

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price for all customers and separately by customer class. Within each pair of columns, the sample is split at the median into "High" (above-median) and "Low" (at-or-below-median) subsamples based on the utility-state's pre-sample balancing-authority fossil fuel share, measured over 2005–2007. Panel A reports estimates for the full sample. Panel B reports estimates for utilities operating in restructured states, while Panel C reports estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Impact of Data Center Entry on Customer Advances for Construction

	All	Restructured	Non-Restructured
	(1)	(2)	(3)
PostDC	7.91*	11.96**	5.48
	(4.59)	(5.21)	(7.49)
YMean ($ million)	10.57	5.02	15.71
ATT/YMean (%)	74.83	238.25	34.88
# of Utilities	106	38	68
N	1,590	570	1,020

Notes: The unit of observation is the utility-year, and the sample only includes investor-owned utilities. The table reports the estimated overall average treatment effect on the treated (ATT). The outcome is customer advances for construction (CAC), i.e., the upfront cash contributions customers make toward infrastructure built to serve them, in millions of 2024 dollars. Column (1) reports estimates for the full sample. Column (2) reports estimates for utilities only operating in restructured states, while Column (3) reports estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A16: Descriptive Statistics of Large Load Tariffs by Electricity Industry Structure

	All (1)	Restr. (2)	Non-Restr. (3)	Diff. (4)
Has any large load tariff	0.038	0.029	0.041	-0.012
	(0.006)	(0.011)	(0.007)	(0.014)
Has strong cost-recovery tariff	0.026	0.025	0.026	-0.001
	(0.005)	(0.010)	(0.006)	(0.012)
Tariff explicitly targets data centers	0.018	0.017	0.019	-0.002
	(0.004)	(0.008)	(0.005)	(0.010)
Minimum monthly demand payment	0.031	0.029	0.032	-0.003
	(0.005)	(0.011)	(0.006)	(0.013)
Contract-term / min. demand payment	0.038	0.029	0.041	-0.012
	(0.006)	(0.011)	(0.007)	(0.014)
Exit / early-termination fee	0.023	0.021	0.024	-0.003
	(0.005)	(0.009)	(0.005)	(0.011)
Upfront contribution (CIAC)	0.013	0.008	0.015	-0.007
	(0.004)	(0.006)	(0.004)	(0.008)
Observations	1,048	241	807	1,048

Notes: The unit of observation is the utility-state. Each cell reports the average share of utility-states for which the listed large-load tariff provision is ever present. Column (1) reports estimates for the full sample. Columns (2) and (3) report estimates for utility-states in restructured and non-restructured states, respectively. Column (4) reports the difference in means between utility-states in restructured states and those in non-restructured states. In columns (1)–(3), values in parentheses are standard errors of the mean. In column (4), values in parentheses are standard errors from two-sample t-tests. The remaining rows report each provision separately. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A17: Heterogeneity in Data Center Effects on Retail Price by Eventual Use of Large Load Tariffs

Dep. Var.:	All		Residential		Commercial		Industrial	
ln(Avg. Retail Price)	No	Yes	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Full Sample								
PostDC	0.026**	-0.022	0.019*	-0.011	0.027***	-0.015	0.042**	-0.028
	(0.011)	(0.019)	(0.010)	(0.019)	(0.010)	(0.020)	(0.020)	(0.027)
YMean ($/kWh)	0.14	0.12	0.15	0.15	0.14	0.13	0.11	0.09
# of Utilities	939	34	939	34	939	34	939	34
N	13,897	836	13,897	836	13,897	836	13,897	836
B. Restructured States								
PostDC	0.040**	-0.024	0.041**	-0.005	0.031*	-0.021	0.057	-0.076
	(0.019)	(0.034)	(0.017)	(0.029)	(0.018)	(0.041)	(0.035)	(0.058)
YMean ($/kWh)	0.15	0.14	0.17	0.16	0.15	0.13	0.11	0.10
# of Utilities	228	11	228	11	228	11	228	11
N	3,256	165	3,256	165	3,256	165	3,256	165
C. Non-Restructured States								
PostDC	0.006	-0.022	-0.004	-0.013	0.020*	-0.014	0.013	-0.021
	(0.011)	(0.022)	(0.009)	(0.023)	(0.012)	(0.023)	(0.021)	(0.028)
YMean ($/kWh)	0.13	0.12	0.14	0.15	0.14	0.13	0.10	0.09
# of Utilities	725	30	725	30	725	30	725	30
N	10,641	671	10,641	671	10,641	671	10,641	671

Notes: The unit of observation is the utility-state-year. The table reports the estimated overall average treatment effect on the treated (ATT). The dependent variable is the log average retail electricity price for all customers and separately by customer class. Within each pair of columns, the sample is split into subsamples based on whether the utility implements a large-load tariff by 2026. Panel A reports estimates for the full sample. Panel B reports estimates for utilities operating in restructured states, while Panel C reports estimates for utilities operating in non-restructured states. YMean reports the pre-treatment mean outcomes among treated units. Standard errors, clustered at the utility-state level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Effect of Data Center Entry on Average Retail Prices When Controlling for Wholesale Prices

Dep. Var.: ln(Avg. Retail Price)	No Wholesale Controls (1)	Control for LMPs (2)
A. Full Sample		
PostDC	0.038***	0.034***
	(0.012)	(0.012)
YMean ($/kWh)	0.14	0.14
# of Utilities	487	487
N	4,919	4,919
B. Restructured States		
PostDC	0.042***	0.036**
	(0.014)	(0.014)
YMean ($/kWh)	0.15	0.15
# of Utilities	186	186
N	1,801	1,801
C. Non-Restructured States		
PostDC	0.016	0.020
	(0.024)	(0.025)
YMean ($/kWh)	0.14	0.14
# of Utilities	312	312
N	3,118	3,118

Notes: The unit of observation is the utility-state-year. The table reports aggregated average treatment effects on the treated (ATT) estimated using CSDID. The outcome is log average retail electricity price across all customer classes. Panel A uses the full wholesale-matched sample. Panels B and C restrict to restructured and non-restructured states, respectively. Column (1) includes no wholesale price controls for the sample of utilities matched to wholesale market exposure. Column (2) controls for day-ahead locational marginal prices. YMean reports the pre-treatment mean retail price among treated units, in 2024 dollars per kWh. Standard errors in parentheses are clustered at the utility-state level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

References

Callaway, Brantly, and Pedro H C Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis: An Annual Publication of the Methodology Section of the American Political Science Association*, 20(1): 25–46.

Lazar, Jim. 2016. "Electricity Regulation in the US: A Guide. Second Edition." The Regulatory Assistance Project., Montpelier, VT.