Do Data Centers Raise Your Electric Bill? Causal Evidence from the United States

By Asa Watten, John Bistline, and Geoffrey Blanford *

Draft: March 27, 2026

We estimate that data centers caused average retail electricity rates to fall modestly in the contiguous United States from 2019 to 2024, both within and across state boundaries. Despite prevailing sentiment, this finding is economically intuitive due to economies of scale and falling generation technology costs. We show patterns of economies of scale can be found for transmission, distribution, and generation costs as well as within and across retail customer classes. Our results imply that, without substantial supply constraints, durable load growth typically reduces average prices in the long-run, all else equal.

In 2024 data centers used about 4.5% of electricity in the United States, more than doubling the sector's power consumption in the previous 5 years (Electric Power Research Institute, 2026). Driven by a rapid acceleration in artificial intelligence (AI) and other data-intensive services, this has led to push-back from some rate-payers and their representatives in government. State legislatures have introduced bans on the construction of new data centers; numerous cities and towns have either enacted moratoriums or resisted local projects; and opinion poling suggests about a third of Americans are opposed to new data centers near them (Bordelon and Miller, 2026). Concerns about data centers include air pollution,

* Watten: EPRI (awatten@epri.com); Bistline: EPRI; Blanford: EPRI

noise, water consumption, and increasingly, higher utility bills. Perceived impacts on rates are key point of the contention, motivating the "Ratepayer Protection Pledge Proclamation" issued by the White House and signed by 7 companies.

While rate-payers fear new load will increase prices, history supports the opposite hypothesis. In the second half of the 20th century, US electricity demand increased while real prices decreased. In this century, both price and demand trends flattened. Demand grew at an annual rate of less than 1% on average from 2014 to 2024. During this time average retail electricity prices increased by 2.2% per year on average, about keeping pace with inflation (U.S. Energy Information Administration, 2025). 2021 and 2022 were anomalous in nominal terms with the average price in the US increasing by 4% and 11% respectively, though keeping pace with the high-levels inflation in those years. These recent national price trends are driven by localized price changes in the Northeast and West Coast. California in particular saw average prices increase nearly 40% from 2019 to 2024 owing largely to wildfire costs (ibid).

Is this recent heterogeneity in electric bill inflation driven by the location of new data centers? At first glance, there does not appear to be an obvious relationship between data centers and price changes (see Figure 3). Virginia, which is far and away the leader in data centers showed middling rate inflation from 2019 to 2024. However, there are clear economic incentives for investors to locate in states with cheaper electricity in expectation, and this could lead to a correlation between data centers and relatively lower prices. So, we want to know: have data centers caused rates to be higher than they would have been without them?

The argument of this paper is simple. Without substantial constraints on building infrastructure, power systems will typically demonstrate constant or falling incremental costs—assuming a fixed load shape—as total load increases.[1] Thus, average prices fall as fixed-cost recovery is spread across more kilowatt-hours, and

[1] A version of this argument is presented in Electric Power Research Institute (2025).

falls further if incremental costs are declining.[2] Our descriptive and causal results are consistent with this returns to scale argument. The implications of this result are important for other types of load-growth including the electrification of transportation and heating. We discuss the limitations of this argument in the discussion at the end.

Electricity rates within a utility are largely determined by their customer class. Could it be that data centers, which are classified as either industrial or commercial, lower overall system costs while increasing residential rates? Industrial and commercial customers do typically see lower rates compared to residential customers. So, perhaps residential customers are subsidizing the incremental fixed-costs of these lower-rate customers. To address this concern, we estimate these cross-class effects on residential prices.

For the period 2019 to 2024 in the contiguous United States (CONUS), we show: 1) The four largest sub-categories of power systems costs—capital expenditures (capex) and operational expenditures (opex) of both transmission and distribution (T&D) and generation—fall per unit electricity delivered as demand increases on average; 2) We observe evidence of own-class economies of scale and lower prices when other customer classes increase their demand; 3) Turning to data centers, new data center capacity from 2019 to 2024 was not correlated with higher average prices for any customer class—instead, the average data center was built in a state where prices changed with the national average; 4) Using an instrumental variables approach to estimate the causal effect of data centers on prices, we estimate that data centers caused prices to modestly decrease on average; 5) We do not find evidence of price spillovers from neighboring states.

For the instrumental variable, we use the total length of the 1947 Eisenhower

[2]Why would incremental costs fall? Larger scale may allow for higher utilization of the existing generation (Betancourt and Edwards, 1987) distribution network (Kwoka Jr, 2005; Yatchew, 2000), depending on the change in load shape. Greater demand density drives down incremental costs (ibid). Newly built generating assets are cheaper and more efficient than their predecessors (Bolinger, Wiser and O'Shaughnessy, 2022; Way et al., 2022; Mirletz et al., 2024; Cole, Frazier and Augustine, 2023).

Interstate Highway Plan within each state. We argue this instrument is exogenous to electricity prices conditional on GDP and population. Why are instrumental variables necessary? It seems plausible a priori that data center developers would be incentivized to locate in places where future electricity prices are expected to be low even with load growth. This would cause negative bias in the ordinary least squares (OLS) estimate. Alternatively, we can tell a speculative story that leads to bias upwards: utilities that clear interconnection queues faster have higher capacity to update their infrastructure; having more capacity to build is costly; and data centers prefer to locate in places where there is quick access to power, which tends to be places with higher prices. There may be other plausible stories which requires IV approach to adjudicate between hypotheses.

We estimate the elasticity of average residential retail electricity prices with respect to data center capacity additions to be -0.02. That is for a doubling of data center capacity, prices fall by 2% on average (see Table 2). While instrumental variables have well-known size inflation effects, the strength of the instrument is such that it is not a major concern. Further, robust methods clearly reject positive effects of data centers on retail prices ($p<0.001$). We present estimates for other customer classes in the appendix. These estimates are remarkably similar across classes.

Observations in our main analysis are state-years. Because electrical grids cross state boundaries, we should be concerned that the spillover effects of data centers on their neighbors could bias our results. Before testing the effect of spillovers, we assumed they would be price increasing. Our reasoning went: higher demand could bid up the regional wholesale price of electricity while savings from economies of scale would be passed through only to local customers. If this was so, our results would be biased downwards. However, our results do not support price spillovers from neighbors—the estimated elasticity is 0.001 with a standard error of 0.013 (see Table 3).

This study adds to the literature on economies of scale in electricity markets since Christensen and Greene (1976), and contributes to the recent literature on the determinants of retail electricity prices in the US (e.g. Wiser et al. 2025; Borenstein and Bushnell 2022) and on the potential rate effects of data centers (e.g. Knittel, Senga and Wang 2025).

This study is unique in showing causal effects of long-lived changes in electricity demand. While Wiser et al. (2025) shows negative correlation between average prices and changes in electricity sales, the study does not identify causal effects for the reasons state above. Within economics, the studies that estimate causal effects have sofar used identifying variation that is limited to the short-run. For example, Dong et al. (2025) use random variation in annual temperature to estimate the effect of demand changes on electricity prices while Holland et al. (2016) use random variation in hourly demand by controlling for hour-month fixed effects to estimate marginal changes in local air pollution. Both of these estimation strategies rely on random variation that is transient. Essentially these studies are limited to estimates of the short-run operating margin and exclude investments in the system that are driven by shifts in future demand expectations.[3]

The paper proceeds as follows: Section 1 describes the data; section 2 provides descriptive evidence of economies of scale for cost and sectoral demand; section 3 describes our estimation strategy and results; section 4 concludes with a discussion of the implications and limitations of these results.

I. Data

A. Data Used for Utility-Level Analysis

To investigate power system costs, we use data collected by the Federal Energy Regulatory Commission (FERC) Form 1 for CONUS. These data contain utility

[3]This relates to an on-going debate over how to compute marginal emissions for energy demanding or supplying investments. See Bistline and Watten (2025).

level data of costs broken down by T&D and generation. These are further disaggregated into operation and capital expenses. The utilities represented in these data are exclusively large investor-owned utilities. We exclude utilities with any years with missing or no residential, industrial, or commercial customers, which removes wires-only utilities. Total sales of the remaining 366 utilities in the sample range from 3 Gigawatt-hours to 162 Terawatt-hours. These data account for 69% of retail electricity sales.

Importantly, FERC exempts some wholesale generators from reporting costs. Therefore, some generation costs are missing or show up exclusively in the generation opex of the utility buying the wholesale power. The extent to which this is an issue is discussed in Section II.B.

B. Data Used for State-Level and Causal Analysis

Retail revenue and sales come from US Energy Information Administration (EIA) Form 861 as cleaned and compiled by Selvans et al. (2026). Generator capacity additions for solar, wind, natural gas, and storage are from EIA Form 860.

To measure data center expansion, we use data from S&P Global Market Intelligence (2026), which are compiled from various sources including public permitting records, site visits, briefings and survey. The data include the year in which data centers became operational or were expanded, along with nominal IT capacity of each facility. Capacity information is missing for some older facilities, which leads us to restrict our analysis to the six-year period from 2019 to 2024. This period coincides with the rapid growth of hyper-scaler operators. Over these six years, these data show between 1.6 and 6.2 gigawatts of capacity were added annually to CONUS; 17.2 gigawatts in total.

Annual state level population estimates are from U.S. Census Bureau (2020, 2024) while GDP estimates are from U.S. Bureau of Economic Analysis (2025).

The 1947 Dwight D. Eisenhower National System of Interstate and Defense Highways Plan geographic data were downloaded from the National Bureau of Economic Research website as used in Brinkman and Lin (2024).

II. Costs and Prices at Scale

This section formalizes a model of power system costs and average prices. It then uses insights from the model to interpret descriptive statistics of both utility cost data and the relationship between customer-class demand and customer-class retail prices.

A. What Determines Prices?

In retail markets for electricity, price does not equal marginal cost (Borenstein and Bushnell, 2022). Instead, customers are charged varying combinations of fixed and volumetric charges, which are expected to recover the utilities costs and allow for some profit (depending on the utility structure). Utilities may also offer tiered pricing, time of use pricing, or special pricing for specific uses such as electric vehicle charging.

So, it is standard practice to use average all-in prices to make prices comparable between utilities (e.g. Dong et al. 2025; Borenstein and Bushnell 2022; Holland et al. 2016). Average prices $\bar{p}_t$ are simply total revenue TR_t divided by total sales q_t:

$$\bar{p}_t = TR_t/q_t. \tag{1}$$

We can disaggregate total revenue into the components of total cost

$$
\begin{aligned}
(2) \qquad TR_t =& TC_t + \text{markup} \\
=& FC_t^{TD}\big(\Delta A^{TD}(\mathbf{E}_t[q_t, q_{t+1}, ...]|A_{t-1}^{TD}), A_{t-1}^{TD}\big) \\
&+ VC_t^{TD}\big(q_t|A^{TD}\big) \\
&+ FC^{G}\big(\Delta A^{TD}(\mathbf{E}_t q), A_{t-1}^{G}\big) \\
&+ VC_t^{G}\big(q_t|A^{G}\big) \\
&+ \text{markup}
\end{aligned}
$$

where: TC_t is total cost in time t, which can be disaggregated; FC are fixed costs; VC are variable costs; A are assets; TD superscripts indicate T&D while G indicates generation; q is demand for power in the appropriate units e.g. terawatt-hours; Δ is the difference operator with the previous period while E is the expected value operator.

The cost-minimizing set of new capital investments $\Delta_t A^x$ is a function of the stream of current period and future demand q given the set of existing assets A_{t-1}^x. Fixed costs are functions of the capital expenditures of new assets and the amortized cost of capital for existing assets. Variable costs are functions of demand in the current period q_t, owing to upward-sloping short-run marginal abatement costs, which determines dispatch order (to a first approximation) in many markets. We can assume that each function is at least weakly increasing in q. For simplicity, we omit weather and other shocks that may damage assets, the retirement of existing assets, and administrative costs. We also omit time of day, week, and year considerations, which we discuss later.

To describe a durable shift in demand, let $\tilde{q}$ be a fixed change in current and future demand—i.e. $\{q_t + \tilde{q}, q_{t+1} + \tilde{q}, ..., q_{t+n} + \tilde{q}\}$. What happens to average

prices when $\tilde{q}$ changes? To find the direction of the price change, we can sign

$$\frac{dTR_t/q}{d\tilde{q}} = \frac{dFC^{TD}/q}{d\tilde{q}} + \frac{dVC^{TD}/q}{d\tilde{q}} + \frac{dFC^{G}/q}{d\tilde{q}} + \frac{dVC^{G}/q}{d\tilde{q}} + \frac{d\text{markup}/q}{d\tilde{q}}. \tag{3}$$

But, what is the sign of each term? Markups can be ignored, assuming that they are either additive or multiplicative. It can be shown that if one of the components of total costs is convex (concave) then the corresponding term is positive (negative).[4] Due to economies of scale and falling technology costs, we expect the fixed costs to be concave and the corresponding terms to be negative. The shape of variable cost curves is more ambiguous.

Regarding the variable cost of generation, we know from observing convex dispatch curves that the variable cost of generation for a fixed set of assets is convex. Note, however, that $\frac{dVC^G}{dq} = \frac{\partial VC^G}{\partial q} + \frac{\partial VC^G}{\partial A^G}\frac{\partial A^G}{\partial q}$. That is, in the long-run, variable costs do not only shift along the dispatch curve ($\frac{\partial VC^G}{\partial q}$), which we know to be convex, but the dispatch curve transforms with new capital investment ($\frac{\partial VC^G}{\partial A^G}\frac{\partial A^G}{\partial q}$). We can also assume that this second term is negative or the investment would not be made. However, we do not know which term dominates to sign the total long-run effect. Note that studies that use short-run random variation to identify costs are only able to capture the dispatch effect $\frac{\partial VC^G}{\partial q}$ ("operating margin") and completely ignore investment effect ("build margin").

We can make the same argument, though it is more speculative, for the variable T&D cost term. Given existing infrastructure there may be convex costs as increased load creates more capacity constraints in the network, though this need not be so. As with generation, we expect that transmission investments in the long-run would ameliorate this.

Returning to fixed costs, if T&D can expand quantity with relatively small

[4]Proof sketch for the case $f(q)$ strictly concave. Assume $q \geq 0$, and $f(0) \geq 0$ (costs are never negative). Let $g(q) = f(q)/q$. We want to show $f(q)$ strictly concave implies $g'(q) < 0$. $g'(q) = [f'(q)q - f(q)]/q^2$ implies $\text{sign}(g'(q)) = \text{sign}(f'(q)qf(q))$. $f(q)$ strictly concave implies $f(y) < f(x)+f'(x)(y-x)$. Therefore, $0 \geq f(0) < f(x) - f'(x)x$. So, $0 \leq f(0) \leq f(q) - f'(q)x$ and $f'(q)q - f(q) < 0$ and $g'(q) < 0$. The convexity case follows directly.

changes in fixed costs (e.g. using the same land or poles), then we expect the T&D fixed cost term ($\frac{dFC^{TD}/q}{d\tilde{q}}$) to be negative. In genera, we expect this to be true when new load is located in areas where service is already provided.

Finally, we expect the generation fixed cost term ($\frac{dTC^{G}/q}{d\tilde{q}}$) to be negative due to falling technology costs. This may not be the case if new load has very different load shapes compared to the past or in the presence of escalating costs due to supply chain constraints or permitting bottlenecks.

B. Descriptive Evidence of Economies of Scale by Cost Type

Ultimately, the sign of the effect of durable demand changes on average prices is an empirical question. Using our panel of large investor-owned utilities, we estimate the correlation of sales and four cost categories: operational expense (variable cost) and capital expense (fixed cost) for generation and T&D.

Overall the data show a pattern consistent with economies of scale in all sub-categories of cost. Figure 1 displays the logged data for electricity generation capital expenses and operational expenses, while figure 2 shows the same for T&D. A 1% increase in electricity sales is correlated with a 1.3% increase in total generation capital expenses or a 0.4% increase in average costs (dollars per megawatt-hour). This would seem to show a convex relationship and violate economies of scale. However, when controlling for year and utility effects (the log de-meaned by year and utility), the relationship becomes negative: a 1% increase in sales is associated with a 13% drop in capital expenses. Similarly, T&D capital expenses appear to be convex in sales; a 1% increase in sales is associated with a 1% increase in costs or a 0.2% increase in average costs. However, once year and utility effects are removed, the relationship is strongly negative: a 1% increase in sales is associated with a 2% decrease in average costs.

On the operations side, the picture is even clearer. Comparing between utilities (no utility fixed effects) and within utilities (utility fixed effects) average oper-

ational expenses are negatively correlated with sales; a 1% increase in sales is associated with a 2% decline in average cost when comparing between utilities, and a 16% decline comparing within utilities for electricity generation and 1.4% and 8% declines respectively for T&D.

As noted above, FERC exempts wholesale generators from reporting. For utilities that purchase power from the wholesale market, this could show up as shifting generation capex to generation opex. This may lead to biased estimates of the relationship for the individual cost category. As the signs of both opex and capex regressions are negative when utility fixed effects are included, we can be confident that the overall relationship is negative within this sample of utilities.

We may be concerned that while the estimated slope is consistent with economies of scale on average, this relationship may switch at some point. Visually, we see no evidence for this–there are no obvious swings upwards in any of the figures–and will leave non-parametric estimates of the relationship for future work.

C. Descriptive Evidence of Economies of Scale by Customer Class

We now investigate the relationship between sectoral demand and own- and cross-sector prices and switch to using state-level data for CONUS, which is includes data for the missing 31% of electricity sales in FERC Form 1 data. Table 1 shows the estimated correlation between log demand by customer-class and customer-class revenue. The first three columns show results for log residential revenue as the dependent variable, while the last three show log industrial revenue as the dependent variable. Columns 1 and 4 make comparisons across states using year fixed effects while columns 2 and 5 compare changes within states by using both year and state fixed effects. Columns 3 and 6 include demand from neighboring states. Note that coefficients less than 1 indicate a concave correlation and therefore an average price decreasing relationship.

Both across and within states, the own-sector relationship is estimated to be

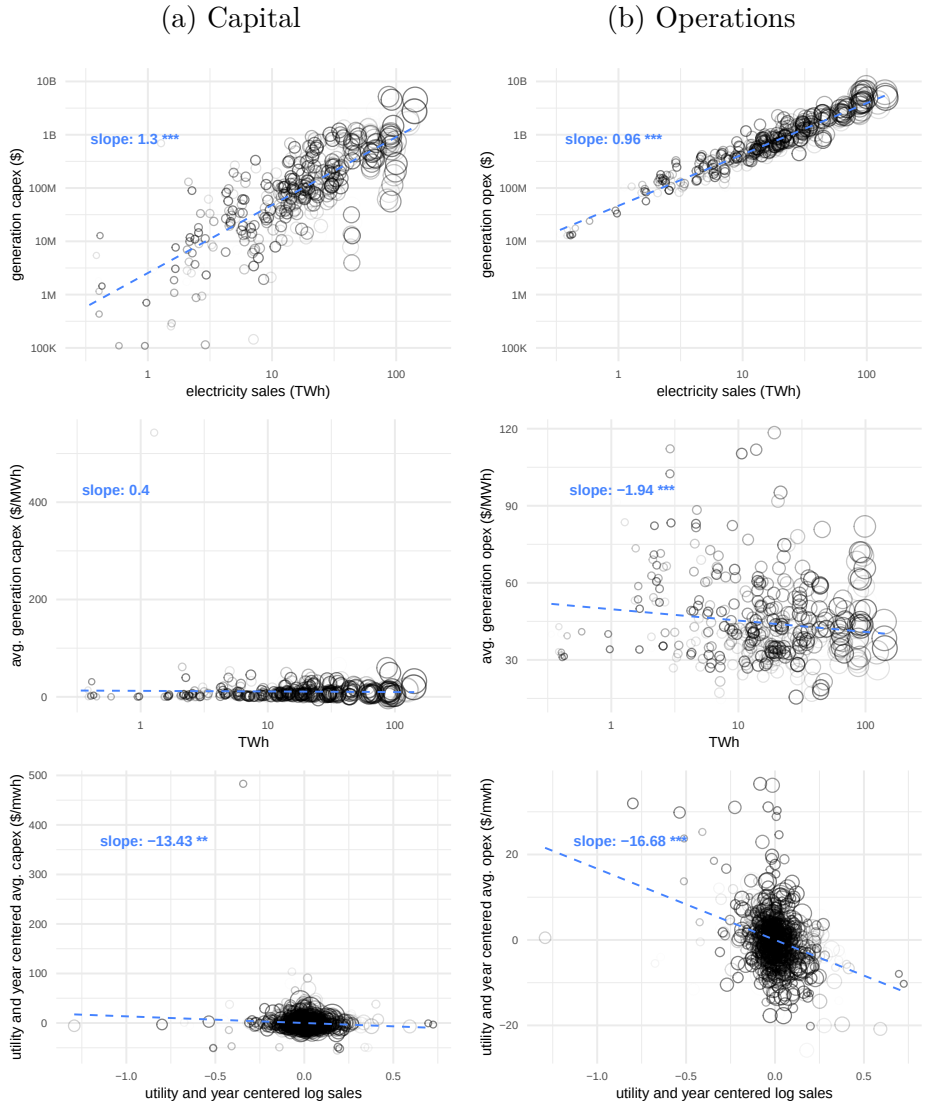


Figure 1. : Generation Annual Costs

Note: Scales are not fixed across panels. The scale of the top four panels have horizontal axes in logs with labels in levels. The bottom two panels are de-meaned by both utility and year and the horizontal axes are log differences. Data spans from 2019 (lightest) to 2024 (darkest). Bubble size is scaled to total electricity sales. Data are from Federal Energy Regulatory Commission Form 1. Slopes are estimated via OLS. Regression tables for each panel are provided in the Supplementary Appendix.

concave (average-price decreasing). Additionally, both across and within states, industrial and residential demand have a negative relationship with the revenue

(a) Capital (b) Operations

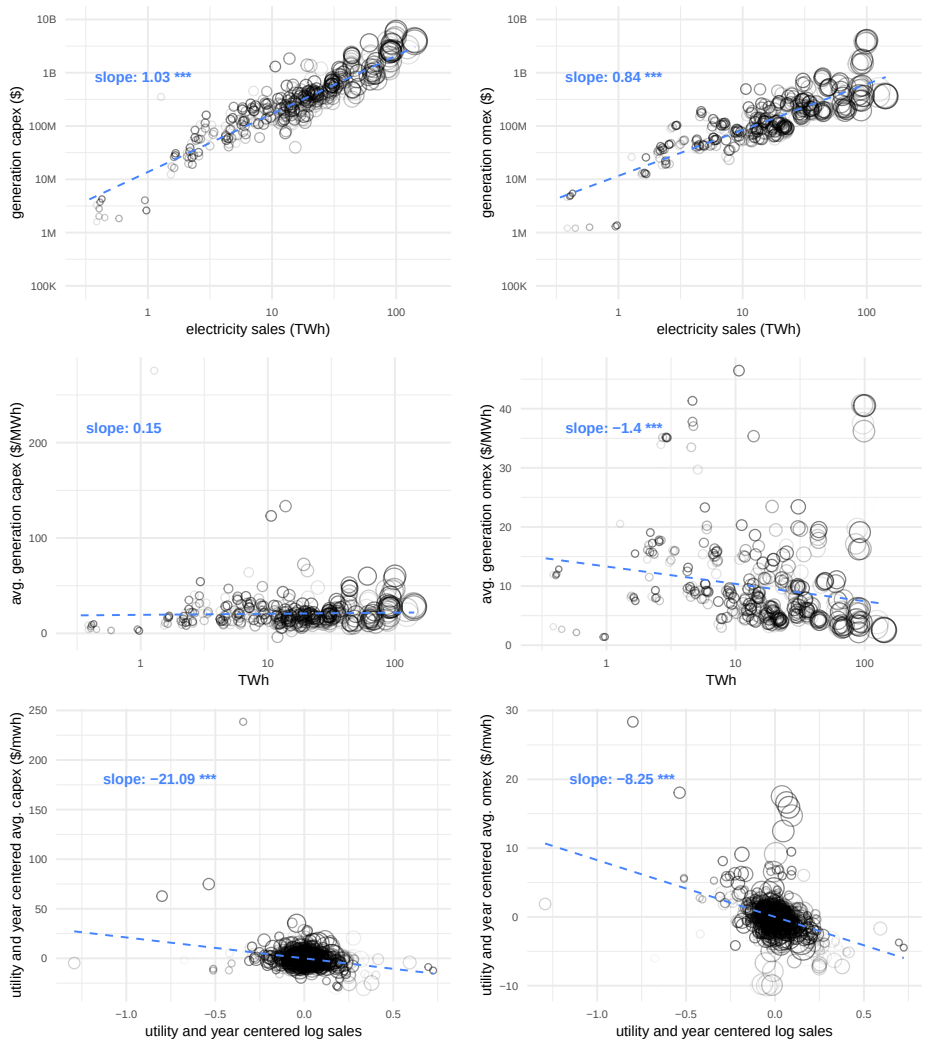


Figure 2. : Transmission and Distribution Annual Costs

Note: Scales are not fixed across panels. The scale of the top four panels have horizontal axes in logs with labels in levels. The bottom two panels are de-meaned by both utility and year and the horizontal axes are log differences. Data spans from 2019 (lightest) to 2024 (darkest). Bubble size is scaled to total electricity sales. Data are from Federal Energy Regulatory Commission Form 1. Slopes are estimated via OLS. Regression tables for each panel are provided in the Supplementary Appendix.

of the other. Commercial demand shows a positive relationship with revenues when comparing across states but a negative one when comparing within states.

Table 1—: Cross-Sector Price and Demand Correlation

	Log residential rev.			Log industrial rev.		
	Between	Within	Spills	Between	Within	Spills
Log industrial sales	-0.170***	-0.231***	-0.222***	0.882***	0.515***	0.527***
	(0.016)	(0.059)	(0.060)	(0.017)	(0.089)	(0.089)
Log residential sales	0.940***	0.504***	0.571***	-0.127*	-0.088	-0.098
	(0.048)	(0.124)	(0.156)	(0.051)	(0.186)	(0.230)
Log commercial sales	0.166***	-0.147**	-0.136*	0.183***	-0.057	-0.029
	(0.045)	(0.054)	(0.055)	(0.049)	(0.081)	(0.082)
Log industrial sales of neighbors			-0.140			-0.529
			(0.220)			(0.325)
Log residential sales of neighbors			-0.233			-0.840
			(0.640)			(0.946)
Log commercial sales of neighbors			-0.135			1.022**
			(0.235)			(0.348)
Year fixed eff.	X	X	X	X	X	X
State fixed eff.		X	X		X	X
R^2	0.967	0.998	0.998	0.963	0.995	0.995
Observations	288	288	288	288	288	288

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
The dependent variable is given by column-group title. Observations are state-years. "Between" estimates make comparisons across states while controlling for common shocks and trends; "Within" uses the variation within states by adding state fixed-effects. "Spills" adds variables for adjoining state neighbors.

These estimates both fit the pattern of economies of scale and of no systematic cost shifting from one sector to another. The point estimate of the correlation with neighboring state demand is generally negative, though imperfectly estimated. The exception is commercial demand from neighbors, which is positively correlated and significant for industrial revenue. We do not have a clear story for interpreting this result.

(a) Residential (b) Industrial

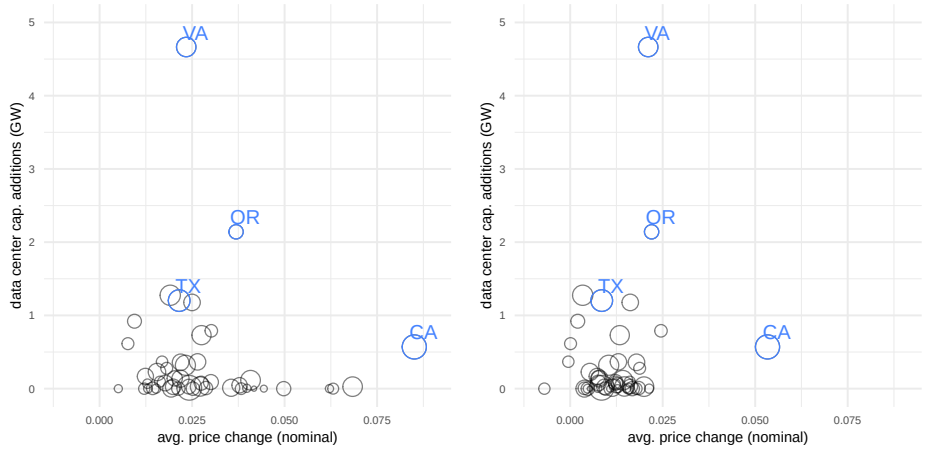


Figure 3. : Average Price vs Data Center Capacity Change

Note: Changes are the difference between 2024 and 2019. Price data are from EIA and data center data are from S&P Global.

III. Estimating the Effect of Data Centers on Average Retail Prices

We now turn to causal evidence for one source of new large load that is of particular concern: data centers. Figure 3 shows the raw data for data center capacity additions compared to nominal average price changes by state for the study period, broken out into residential and industrial prices. Of note is that the state with the highest prices change (California) did not experience the most data center expansion, while the state with the most data center growth (Virginia) is middling in its residential price change–it is on the higher end of industrial price changes. In general there is no apparent correlation between data centers and average price changes.

We estimate aggregate revenue elasticities of the following form

$$\ln Rev^i_{st} = \beta_1 \ln TWh^i_{st} + \beta_2 \sinh^{-1} DC_{st} + \gamma' X_{st} + \alpha_t + \delta_s + \varepsilon_{st} \tag{4}$$

where: Rev^i_{st} is revenue for customer class i in state s and year t; TWh_{ist} is

retail sales; DC_{st} is data center capacity additions since 2018 while $\sinh^{-1}$ is the inverse hyperbolic sine operator[5]; β_2 is our coefficient of interest and is interpreted as the cross-price elasticity of data centers; X_{st} is a vector of state controls; α_t is a year effect that is constant across all states; δ_s are state-specific effect that is constant across all time periods. There are three customer classes i: residential, commercial, and industrial.[6] Only one customer class is included in the dependent variable: the results presented in the body of the paper are all for residential customers. Data centers are either classified as industrial or commercial customers, depending on the state, which may also change over time. Therefore, commercial and industrial sales are omitted as controls as they would absorb the effect of data centers.

Note that this equation estimates revenue for fixed demand for sector i conditional on that sector's demand. Mechanically, coefficients are the all-in average price elasticities for a fixed level of demand. Intuitively, this is the closest measure to what we want to know: if the sector's demand for electricity increases, how much is a customer's bill for a fixed amount of electricity changing? And not: if a customer buys more electricity, how much is their average price changing? Importantly, we do not know the rate structure for any given (or group of) rate payer(s).

α_t controls for shocks and trends that are common to all states—e.g. the COVID-19 pandemic, economy-wide supply chain disruptions, and fuel prices linked to global or national markets. δ_s controls for persistent differences in the characteristics of states (e.g. climate) and aspects population and economy that are fixed for the period. The vector of time-varying controls X_{st} includes the log of population and log of GDP to capture changes to a state's electricity customers

[5]The inverse hyperbolic sine behaves similarly to the natural log, allowing us to estimate elasticities, but, unlike the natural log, allows for zeros in the data, which is necessary in our case. The scale of the variable matters for estimating elasticities (Bellemare and Wichman, 2020). If the mean of the variable is too small, it can cause bias in the interpretation of the estimate. In our case, the mean of capacity additions measured in kilowatts is far above the threshold where bias is rounding error.

[6]Estimating industrial demand requires that we switch to log of average industrial price on the left hand side as we can not control for industrial demand while estimating the effect.

that might effect prices.

A. *Identification*

We estimate regression 4 using ordinary least squares (OLS) and two-stage least squares (2SLS) using an instrumental variable (IV). OLS estimates are potentially biased if uncontrolled-for drivers of utility sales revenue correlate with data center capacity. Of particular concern is how the data center site selection process may be correlated with expectations of lower future prices. Indeed, for any profit-maximizing investor, electricity rates ought to influence site selection. It is also possible that "speed to power" (the length of time for new facilities to be connected to the grid) may be correlated with higher costs. The first type of endogeneity would bias our estimate downward, while the second would bias our estimate upwards. From our conversations with data center developers, "speed-to-power" is often the dominant factor in siting decisions.

To the extent that these siting incentives are constant through our study period, the state fixed effects absorb their influence. Data centers could still be endogenous to average prices if the incentives to locate in one state over another are shifting over time. We expect this to be the case for two reasons: 1) the increased pace of data center additions has put new strain on the interconnection queue which would reveal or make relevant differences in utility interconnection ability; 2) investments to harden the grid for the warmer climate (e.g. wildfire hardening in California) is recent and expensive, which could greatly change future price expectations. These developing phenomena are likely to shift price and "speed-to-power" expectations even over our six year period.

To overcome this source of endogeneity, we use the total length of the original interstate highway plan as an instrumental variable for data centers. Arguably, the 1947 interstate highway plan does not affect electricity prices except by way of their lasting influence on population and economic activity. So, by controlling

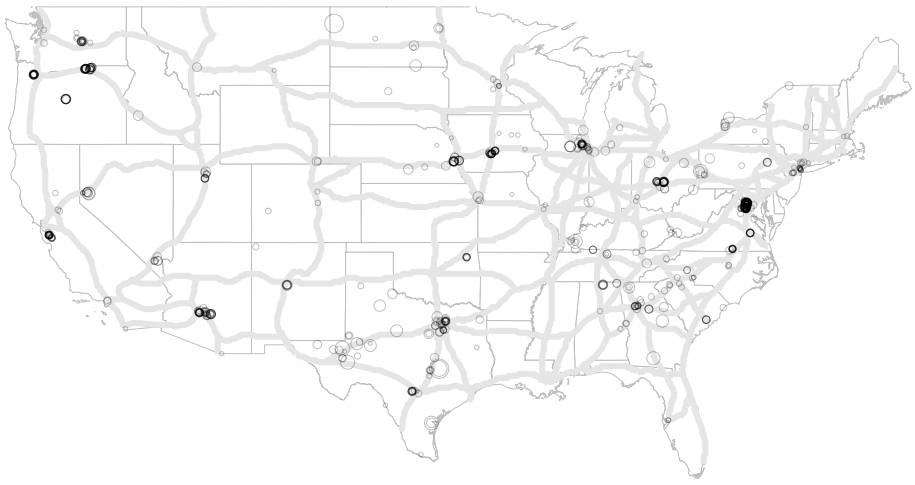


Figure 4. : Data Center Additions and the 1947 Interstate Highway Plan

Note: Circles indicate data centers built between 2019 to 2024. The size of the circles is scaled to data center capacity. The thick gray lines show the original 1947 Dwight D. Eisenhower National System of Interstate and Defense Highways Plan.

for population and state-level GDP, we have cut off these endogenous pathways, thus satisfying the exclusion restriction for instrumental variables. The relevancy of the instrument comes from the observation that terrestrial fiber-optic infrastructure is typically located along interstate highway corridors. This is clearly visible in figure 4, which shows the original 1947 Dwight D. Eisenhower National System of Interstate and Defense Highways Plan overlain with the location of new or expanded data centers during the study period. Further, our instrument is positively correlated with data center capacity additions conditional on the controls ($p < 0.001$). A regression table for the first stage is provided in the Appendix.

Using IV, the first stage is then:

$$\sinh^{-1} DC_{st} = \lambda_1 \ln TWh_{st} + \lambda_2 z_s + \theta' X + \zeta_t + \epsilon_{st}. \tag{5}$$

For the second stage, equation 4 becomes:

$$\ln Rev^i_{st} = \beta_1 \ln TWh^i_{st} + \beta_2 \widehat{\sinh^{-1} DC_{st}} + \gamma' X_{st} + \alpha_t + \varepsilon_{st} \tag{6}$$

where: z_s is the instrument and $\widehat{\sinh^{-1} DC_{st}}$ is the estimated quantity from the first stage. Note that state fixed effects δ_s are now omitted from the second stage. This is because the instrument is time-invariant and therefore is collinear with state fixed effects. This creates a tradeoff. By using the instrument, we must rely more heavily on our controls.

B. Results

Table 2 presents our main OLS and IV estimation results for average residential prices. The first three columns present OLS results. The "between" column omits state fixed effects, allowing for a comparison between states and allowing for selection effects, while the "within state" column includes state fixed effects, using changes within states to estimate the effect, while "within region" swaps state for regional fixed effects.[7] The last two columns present the IV results for the state and state plus region fixed effects estimator. State fixed effects are not used with IV because they are collinear with the instrument.[8] All standard errors are robust to heteroskedasticity and are clustered by state to account for serial correlation. Weak instrumental variables are known to cause size inflation. While both IV specifications have a first-stage F-statistics that are larger than 10 (a commonly used though imperfect critical value), we compute Anderson-Rubin (Anderson and Rubin, 1949) p-values and 95% confidence intervals, which are

[7]Region definitions are collections of states taken from the 2025 U.S. Regional Economy, Greenhouse Gas, and Energy Model (US-REGEN) energy-economy quantitative equilibrium model developed and maintained by EPRI. The 16 regions in US-REGEN are Pacific: WA, OR; California: CA; Mountain-S: NV, UT, NM, AZ; Mountain-N: ID, MT, CO, WY; MISO-N: ND, MN, IA, WI; SPP: SD, NE, KS, OK; MISO-S: LA, AR, MO; Southeast: MS, AL, GA, TN; Florida: FL; South Atlantic: SC, NC, VA; Ohio Valley: OH, KY, WV; Mid Atlantic: PA, NJ, MD, DE, DC; New York: NY; North East: ME, NH, VT, MA, CT, RI; Texas: TX; MISO-E: IL, MI, IN. These regions roughly correspond to North American Electric Reliability Corp (NERC) Reliability Assessment Areas.

[8]Alternative specifications using state fixed effects and the instrument interacted with year or national data center capacity are far to week.

robust to weak-instruments.[9] Our preferred specification is the fifth and final column.

Table 2—: Data Center Average Residential Price Effects

	OLS			IV	
	Between	Within state	Within region	Between	Within region
$\sinh^{-1}$ data center cap.	-0.012**	-0.003	-0.001	-0.072**	-0.018**
	(0.004)	(0.002)	(0.003)	(0.022)	(0.007)
Log residential sales	0.478***	0.317+	0.730***	0.504***	0.633***
	(0.095)	(0.190)	(0.129)	(0.064)	(0.072)
Log population	0.597***	0.008	0.425***	0.808***	0.515***
	(0.145)	(0.657)	(0.120)	(0.150)	(0.072)
Log GDP	-0.080	-0.034	-0.156*	-0.131	-0.101*
	(0.080)	(0.147)	(0.074)	(0.090)	(0.041)
Weak-iv robust 95%CI				[-0.143,-0.046]	[-0.036,-0.007]
Weak-iv robust p-value				0.000	0.001
Year fixed eff.	X	X	X	X	X
State fixed eff.		X			
Region fixed eff.			X		X
First-stage F				14.685	28.501
R^2	0.975	0.998	0.992	0.917	0.989
Observations	240	240	240	240	240

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
The dependent variable for all columns is log residential revenue. An observation is a state-year. As all regressions are conditional on log residential sales, the effects are interpreted as fractional changes in average residential price. All standard errors are robust to heteroskedasticity and clustered by state. The instrumental variable regressions include the Anderson-Rubin 95% confidence interval and p-value, which is fully robust to weak instruments.

These results suggest that data centers did not increase residential rates on average between 2019 and 2024. The OLS estimates are all negative, though the "within state" and "within region" estimates are not statistically different than

[9]Keane and Neal (2024) goes so far as to recommend Anderson-Rubin confidence intervals be used instead of t-tests even with strong instruments.

zero and their magnitudes are quite small (-0.003 and -0.001 respectively). If developers only sited data centers based on static attributes of states, state fixed-effects would be sufficient for estimating a causal effect, and we would conclude that we could not reject no effect. The "within region" OLS estimate is included for comparison to "within state." The similarity of the estimates is promising, indicating that region fixed-effects are capturing much of the same variation.

Turning to the IV estimates, both "between" and "within" estimates are negative and significant at the 1% level. Regional fixed-effects appear to be important as they cause the magnitude of the coefficient of interest to drop by three quarters. The evidence for selection bias from comparing the IV results to OLS is quite subtle. Comparing "within region" IV to "between" OLS, the similarity of the estimates suggests that there is minimal selection on retail rates. Comparing the same to "within region" OLS shows, the effect is much stronger for the causal IV estimate, suggesting modest selection for regions where prices would have increased relative to other regions (dynamic vs static selection), if not for data centers.

In our preferred specification "within region" IV, the point estimate shows that each doubling of data center capacity causes prices to fall by 2%. While the effect is small, we can say with confidence that there is a price reducing effect. The corresponding Anderson-Rubin 95% confidence interval spans -4% to -1% for each doubling and the Anderson-Rubin p-value is 0.001, leading us to reject a 0 null-hypothesis at the 99.9% level. The first-stage F-Statistic (28) is above the threshold where we should be concerned about substantial IV size inflation.[10]

The Appendix shows results for all-sector average prices and industrial average prices. These results are remarkably similar to those presented here for average residential prices. Additionally, the Appendix contains specification charts for these regressions and additional robustness checks. These robustness checks

[10]Keane and Neal (2024) caution "for first-stage F in the 10–20 range [...] the chance that 2SLS will generate an estimate closer to the truth than OLS is remarkably small, unless endogeneity is very severe."

include regressions where California or 2020 are dropped from the sample and specifications without population and GDP controls. These robustness checks yield similar estimates. Additionally, we add controls for new generation by technology (wind, solar, natural gas, and storage). Again the estimate is quite similar, leading us to conclude that transmission efficiencies are pivotal to the price reducing effect of data centers.

Because power transmission crosses state boundaries, we are concerned that the price effects of data centers could spill over to other states and bias results. We know from Hausman (2025); Ham, Kay and Hausman (2025) that there are substantial grid constraints such that generators in neighboring states are more likely to be dispatched in response to higher load than generators further away. So, any spillovers from data centers ought be more concentrated regionally. To test for spillovers, we estimate the effect of data centers in neighboring states. These results are shown in Table 3. The first column shows OLS results with year fixed effects, the second shows OLS results with year and state fixed effects, and the third replaces state fixed effects with regional fixed effects. Instrumental variable estimates are shown in columns 4 and 5 which show year-only and year plus region fixed effects respectively. Own-state sales for all three customer classes are controlled for as well as the residential and commercial sales of neighbors.

No estimated neighboring state effect is statistically different than zero at standard levels. The estimated magnitude drops with the inclusion of state or regional fixed-effects. Without regional fixed effects, the instrument is quite week (first stage F-statistic = 0.026), which leads to a very imprecise and large estimate. Including regional fixed-effects greatly improves the strength of the instrument (first stage F-statistic = 29). The Anderson-Rubin 95% confidence interval for our preferred specification in column 5 spans -0.017 to 0.037 and we fail to reject the null hypothesis of no effect.

Table 3—: Data Center Spillovers

	OLS			IV	
	Between	Within state	Within region	Between	Within region
$\sinh^{-1}$ data center cap. of neighbors	-0.015	-0.006	0.008	0.364	0.010
	(0.011)	(0.004)	(0.008)	(0.794)	(0.013)
Log residential sales	0.675***	0.630*	0.746***	0.432	0.742***
	(0.103)	(0.319)	(0.091)	(0.551)	(0.056)
Log commercial sales	0.008	-0.257**	-0.076	0.005	-0.072+
	(0.082)	(0.081)	(0.056)	(0.156)	(0.038)
Log industrial sales	-0.070**	-0.191+	-0.059*	-0.086	-0.057**
	(0.025)	(0.098)	(0.027)	(0.072)	(0.019)
Log population	0.440***	-0.062	0.486***	1.222	0.489***
	(0.122)	(0.604)	(0.080)	(1.512)	(0.046)
Log GDP	-0.085	0.086	-0.134**	-0.652	-0.137***
	(0.059)	(0.150)	(0.050)	(1.102)	(0.028)
Log residential sales of neighbors	-0.209**	-0.746	-0.253***	-1.315	-0.258***
	(0.075)	(1.252)	(0.067)	(2.187)	(0.043)
Log population of neighbors	0.053	0.573	0.062	-1.484	0.049
	(0.196)	(0.701)	(0.136)	(3.387)	(0.111)
Log GDP of neighbors	0.088	-0.476+	0.075	1.415	0.084
	(0.168)	(0.245)	(0.110)	(2.923)	(0.089)
Weak-iv robust 95%CI				[-Inf,0.029]	[-0.017,0.037]
Weak-iv robust p-value				0.016	0.442
Year fixed eff.	X	X	X	X	X
State fixed eff.		X			
Region fixed eff.			X		X
First-stage F				0.262	29.495
R^2	0.985	0.998	0.994	0.770	0.994
Observations	240	240	240	240	240

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

The dependent variable for all columns is log residential revenue. An observation is a state-year. As all regressions are conditional on log residential sales, the effects are interpreted as fractional changes in average residential price. All standard errors are robust to heteroskedasticity and clustered by state. The instrumental variable regressions include the Anderson-Rubin 95% confidence interval and p-value, which is fully robust to weak instruments. "Neighbors" referrs to all states that share a border.

IV. Discussion

In this study we estimated a modest causal impact of data centers on decreasing residential prices. This is supported by descriptive statistics that indicate economies of scale of power systems. Our findings are not without caveats, which we place into three categories: 1) supply constraints, 2) durability of demand growth, and 3) fuel costs.

Presently, several supply constraints are impacting the U.S. electric sector: orders on new on natural gas turbines are backlogged, which is bidding up their price (Anderson, 2025);); persistent supply chains issues limit the availability of transformers and other electrical equipment;high tariffs on photovolatic arrays from China have increased the price of solar in the United States (Gerarden et al., 2025); and aggressive executive action has stalled federal approvals for wind and solar projects while they are litigated in courts. Each of these constraints make new generation capacity investments more expensive than they would be otherwise and possibly more expensive than incumbent supply. To the extent that these constraints are temporary, price increases can be avoided entirely by delaying interconnection or requiring data centers to pay for the inflated incremental cost of grid expansion.

Crucially, the forward application of our results relies on durable increases in demand for electricity. If, as some expect, the requests for future energy supply to data centers is higher than what they will actually consume, then there is the risk of building a power system that is too large. To address this concern, states including Kansas, Michigan, and Delaware, have pursued financial commitments such as minimum bills or long-term contracts from new large loads.

Finally, our results excluded national-level effects of data center electricity demand such as natural gas prices. We simply do not have the identifying variation to do this, though we expect the effect to have been small. A quick back of the envelope calculation shows data center consumption rose to about 2% of US natural

gas consumption in 2024[11] and 0.04% of global natural gas consumption or 0.06% excluding Russia, China, and Iran.[12] The US has been a net exporter of natural gas since 2017, which has co-integrated natural gas prices with international oil and coal markets (Stock and Zaragoza-Watkins, 2024) making the effective market size much larger and the fuel price effect of US data centers much smaller compared to a domestic shut-in market for natural gas.

The implications of this study for electrification are optimistic. Shifting from natural gas and fuel oil heating to heat pumps and from gasoline vehicles to electric vehicles is expected to substantially increase electricity demand (Bistline et al., 2024) while providing cost benefits to consumers (Electric Power Research Institute, 2025). To the extent that our results hold for different end-uses of electricity, we should not be concerned about electrification increasing prices. However, this study has abstracted away from the load shape—the pattern of load demand within a day, week, and year—of data centers due to a lack of direct observation and potential heterogeneity between data center types. Electrifying heating and transportation have their own load shapes, which could be quite different than data centers.

While hopeful, we are left with concerns for the capacity of the system to build quickly enough to accommodate both rapid data center expansion and electrification simultaneously. Effective grid planning must balance the substantial and clear cost savings from electrification (Rapson and Muehlegger, 2023; Anderson and Kirkpatrick, 2024) with the disruptive but uncertain opportunities arising from AI.

[11] About 40% of natural gas in the US is used for electricity according to the EIA and about 4.5% of electricity was used for data centers in 2024 (Electric Power Research Institute, 2026).

[12] In 2023 the US consumed an estimated 22.2% of global natural gas according the IEA while Russia, China, and Iran together consumed 27.7%.

REFERENCES

Anderson, Jared. 2025. "US Gas-Fired Turbine Wait Times as Much as Seven Years; Costs Up Sharply." *S&P Global Energy*.

Anderson, Soren, and A Justin Kirkpatrick. 2024. "Distributional consequences of policies for electric heat conversion." *Journal of the Association of Environmental and Resource Economists*, 11(S1): S9–S40.

Anderson, Theodore W, and Herman Rubin. 1949. "Estimation of the parameters of a single equation in a complete system of stochastic equations." *The Annals of mathematical statistics*, 20(1): 46–63.

Bellemare, Marc F, and Casey J Wichman. 2020. "Elasticities and the inverse hyperbolic sine transformation." *Oxford Bulletin of Economics and Statistics*, 82(1): 50–61.

Betancourt, Roger R, and John HY Edwards. 1987. "Economies of scale and the load factor in electricity generation." *The Review of Economics and Statistics*, 551–556.

Bistline, John ET, and Asa Watten. 2025. "Emissions reductions of rooftop solar are overstated by approaches that inadequately capture substitution effects." *Nature Climate Change*, 1–3.

Bistline, John, Morgan Browning, Julianne DeAngelo, Daniel Huppmann, Ryan Jones, James McFarland, Anahi Molar-Cruz, Steven Rose, and Steven J Davis. 2024. "Uses and limits of national decarbonization scenarios to inform net-zero transitions." *Joule*, 8(10): 2721–2726.

Bolinger, Mark, Ryan Wiser, and Eric O'Shaughnessy. 2022. "Levelized cost-based learning analysis of utility-scale wind and solar in the United States." *Iscience*, 25(6).

Bordelon, Brendan, and Gabby Miller. 2026. "The big split driving the tricky politics of AI data centers." *https://www.politico.com/news/2026/02/06/tech-industry-ai-data-centers-politics-00762348*, Accessed: 2026-03-01.

Borenstein, Severin, and James B Bushnell. 2022. "Do two electricity pricing wrongs make a right? Cost recovery, externalities, and efficiency." *American Economic Journal: Economic Policy*, 14(4): 80–110.

Brinkman, Jeffrey, and Jeffrey Lin. 2024. "Freeway revolts! The quality of life effects of highways." *Review of Economics and Statistics*, 106(5): 1268–1284.

Christensen, Laurits R, and William H Greene. 1976. "Economies of scale in US electric power generation." *Journal of political Economy*, 84(4, Part 1): 655–676.

Cole, Wesley, A Will Frazier, and Chad Augustine. 2023. "Cost projections for utility-scale battery storage: 2023 update." National Renewable Energy Laboratory Golden, CO.

Dong, Zeyang, Jing Liang, Joshua Linn, and Yueming Qiu. 2025. "How Do Residential Greenhouse Gas Mitigation Technologies Affect Electricity Prices and Consumer Welfare?" Resources for the Future Working Paper 25-28. Accessed: 2026-02-25.

Electric Power Research Institute. 2025. "Win-Win Watts: When Can Data Centers, Efficient Electrification, and New Loads Lower Electricity Prices?" Electric Power Research Institute (EPRI) 3002034619, Palo Alto, CA.

Electric Power Research Institute. 2026. "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption." Electric Power Research Institute (EPRI) 3002034696, Palo Alto, CA.

Gerarden, Todd D, Bryan K Bollinger, Kenneth Gillingham, and Daniel Xu. 2025. "Strategic Avoidance and the Welfare Impacts of US Solar Panel Tariffs." National Bureau of Economic Research.

Ham, Dasom, Owen Kay, and Catherine Hausman. 2025. "Power Flows, Part 2: Transmission Lowers US Generation Costs, But Generator Incentives Are Not Aligned." Resources for the Future.

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

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates. 2016. "Are there environmental benefits from driving electric vehicles? The importance of local factors." *American Economic Review*, 106(12): 3700–3729.

Keane, Michael P, and Timothy Neal. 2024. "A practical guide to weak instruments." *Annual Review of Economics*, 16(1): 185–212.

Knittel, Christopher R, Juan Ramon L Senga, and Shen Wang. 2025. "Flexible data centers and the grid: Lower costs, higher emissions?" National Bureau of Economic Research.

Kwoka Jr, John E. 2005. "Electric power distribution: economies of scale, mergers, and restructuring." *Applied Economics*, 37(20): 2373–2386.

Mirletz, Brian, Laura Vimmerstedt, Greg Avery, Ashok Sekar, Dana Stright, Dayo Akindipe, Stuart Cohen, Wesley Cole, Patrick Duffy, Annika Eberle, et al. 2024. "Annual technology baseline: The 2024 electricity update." National Renewable Energy Laboratory (NREL), Golden, CO (United States).

Rapson, David S, and Erich Muehlegger. 2023. "The economics of electric vehicles." *Review of environmental economics and policy*, 17(2): 274–294.

Selvans, Zane A., Christina M. Gosnell, Austen Sharpe, Zach Schira, Katherine Lamb, Dazhong Xia, Ella Belfer, and Kathryn Mazaitis. 2026. "Public Utility Data Liberation Project (PUDL) Data Release." Accessed: 2026-02-25.

S&P Global Market Intelligence. 2026. "Datacenter KnowledgeBase." *S&P Global Market Intelligence*, Acessed: 2025-12-01.

Stock, James H, and Matthew Zaragoza-Watkins. 2024. "The market and climate implications of us lng exports." National Bureau of Economic Research.

U.S. Bureau of Economic Analysis. 2025. "GDP by State." Accessed: 2026-02-25.

U.S. Census Bureau. 2020. "State Population Totals and Components of Change: 2010–2020." Accessed: 2026-02-25.

U.S. Census Bureau. 2024. "State Population Totals and Components of Change: 2020–2024." Accessed: 2026-02-25.

U.S. Energy Information Administration. 2025. "Electric Power Annual." U.S. Energy Information Administration. Tables 2.2 and 2.4. Accessed: 2026-03-01.

Way, Rupert, Matthew C Ives, Penny Mealy, and J Doyne Farmer. 2022. "Empirically grounded technology forecasts and the energy transition." *Joule*, 6(9): 2057–2082.

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(4): 107516.

Yatchew, Adonis. 2000. "Scale economies in electricity distribution: A semiparametric analysis." *Journal of applied Econometrics*, 15(2): 187–210.

Supplemental Appendix

Table A1: Generation Annual Capital Expense

	capex (1)	avg. capex (2)	avg. capex (3)
ln MWh	1.300***	0.400	−13.426**
	(0.034)	(0.469)	(5.057)
Observations	1043	1068	1066
R^2	0.58	0.00	0.15
utility fixed effects			X
year fixed effects			X

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Table A2: Generation Annual Operation Expense

	opex (1)	avg. opex (2)	avg. opex (3)
ln MWh	0.961***	−1.943***	−16.676***
	(0.007)	(0.305)	(1.572)
Observations	1068	1068	1066
R^2	0.95	0.04	0.81
utility fixed effects			X
year fixed effects			X

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Table A3: Transmission and Distribution Annual Capital Expense

	capex (1)	avg. capex (2)	avg. capex (3)
ln MWh	1.034***	0.151	−21.089***
	(0.015)	(0.306)	(2.541)
Observations	1066	1068	1066
R^2	0.82	0.00	0.50
utility fixed effects			X
year fixed effects			X

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Table A4: Transmission and Distribution Annual Operation and Management Expense

	opex (1)	avg. opex (2)	avg. opex (3)
ln MWh	0.843***	−1.403***	−8.254***
	(0.013)	(0.131)	(0.641)
Observations	1068	1068	1066
R^2	0.80	0.10	0.84
utility fixed effects			X
year fixed effects			X

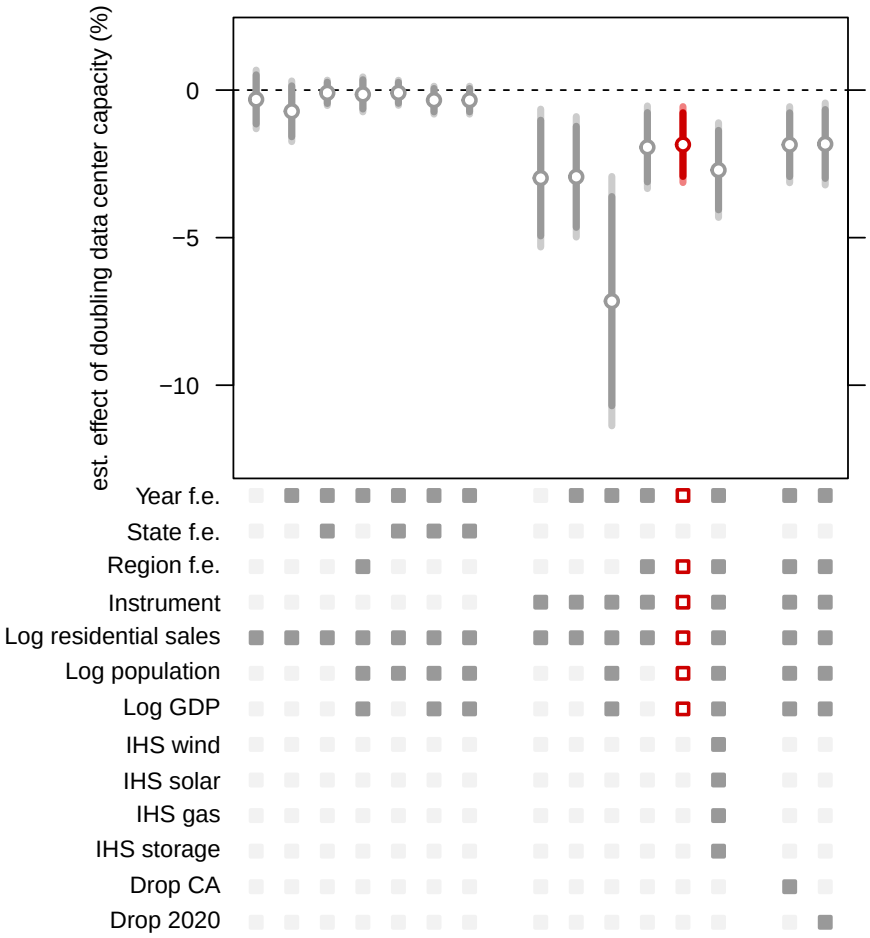
+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Table A5—: First-stage Regression Results

	No controls	Between	Within
Intercept	4.285***		
	(0.486)		
Instrument	3.705***	1.848***	3.357***
	(0.371)	(0.482)	(0.629)
Log residential sales		0.784	-6.678***
		(1.060)	(1.782)
Log population		1.803	4.816*
		(2.051)	(1.960)
Log GDP		-0.326	3.728**
		(1.338)	(1.256)
Year fixed eff.		X	X
Region fixed eff.			X
R^2	0.258	0.393	0.704
Observations	288	240	240

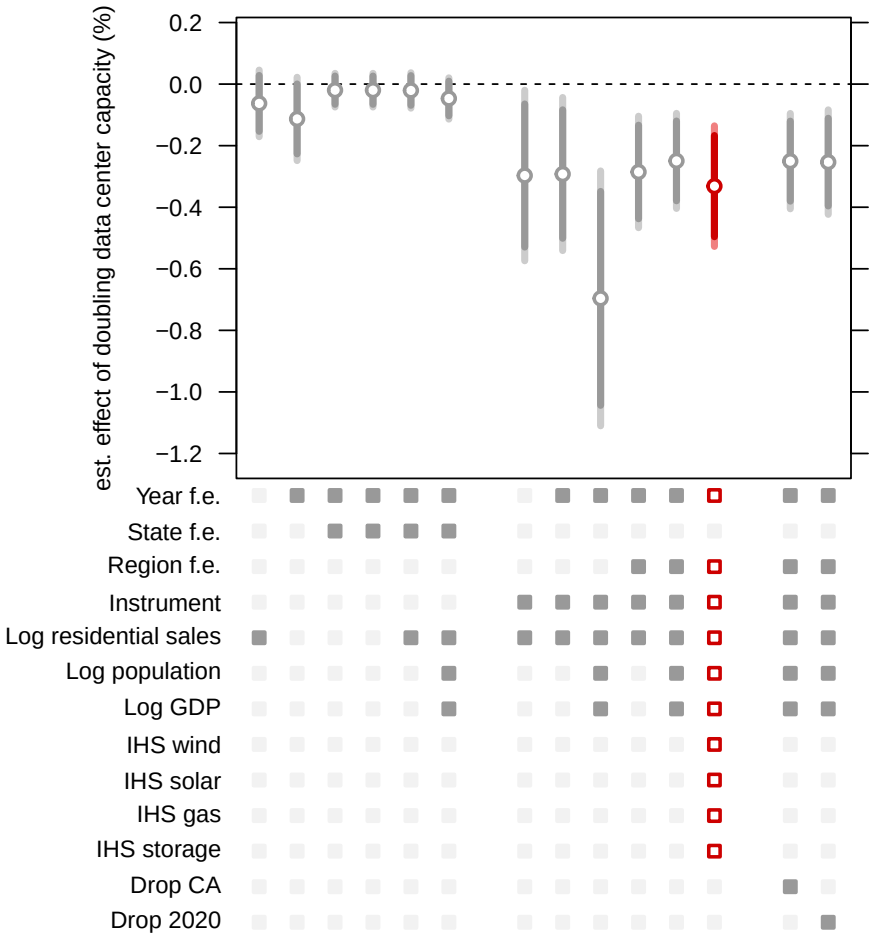
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Figure A1. : Residential Revenue Estimates



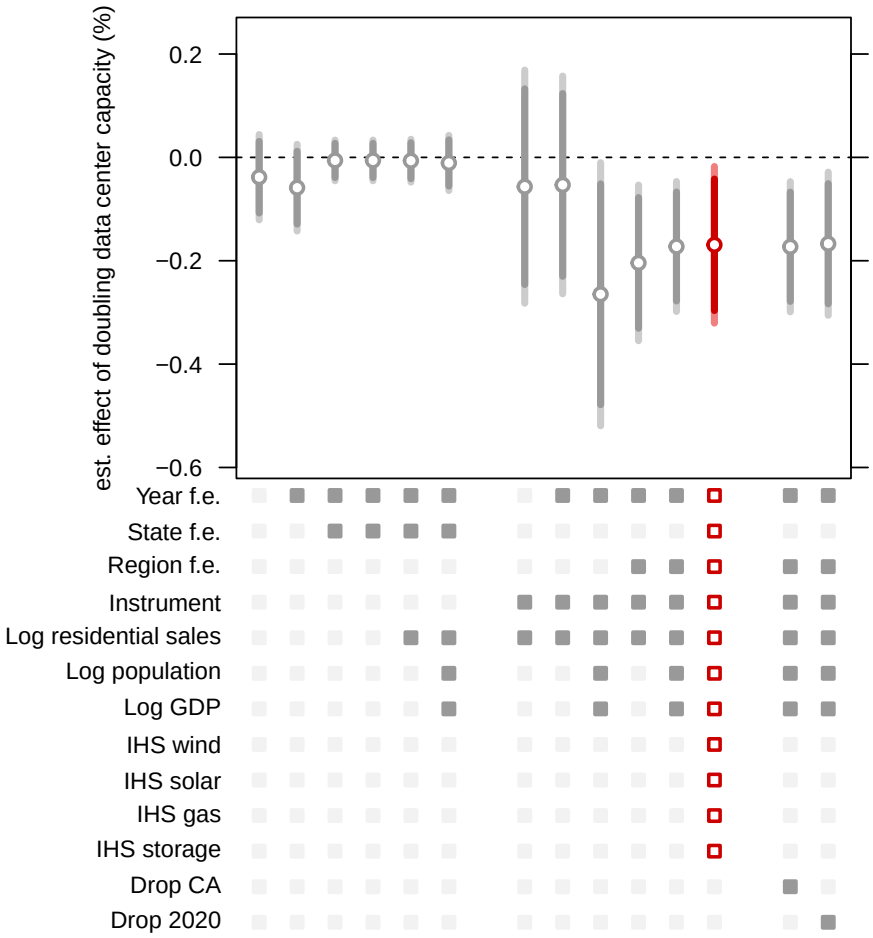
Note:

Figure A2. : All-sector Average Price Estimates



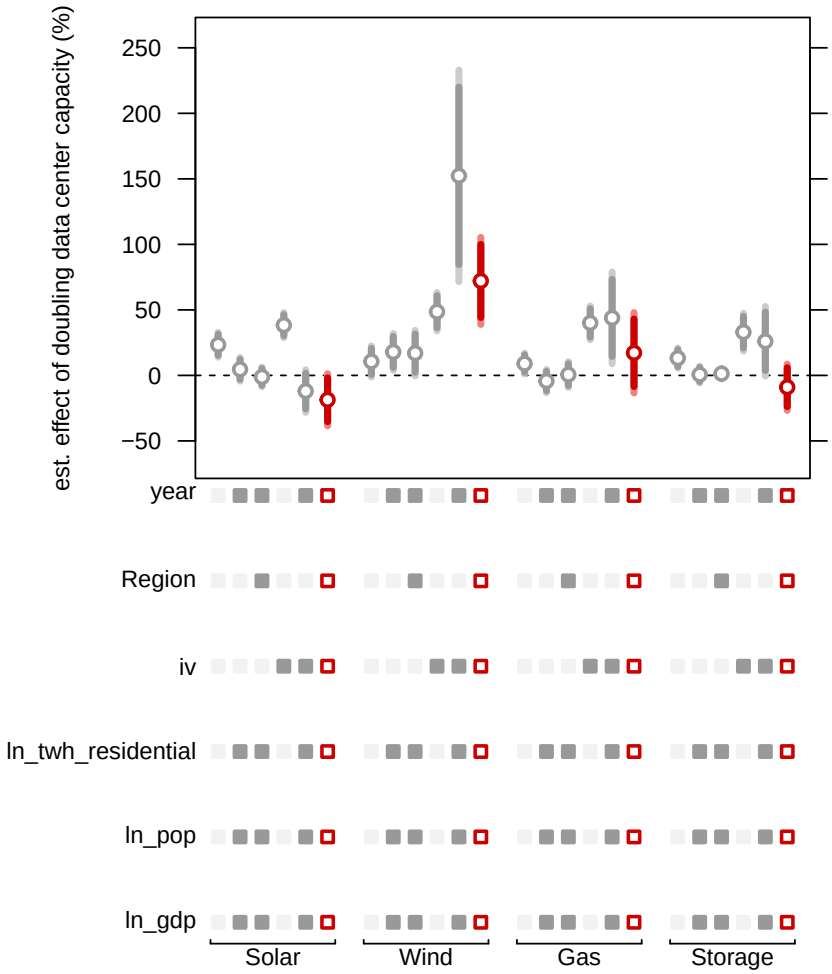
Note:

Figure A3. : Industrial Average Price Estimates



Note:

Figure A4. : New Generation Capacity Estimates



Note: