Data Centers and Local Labor Markets: Evidence from the Texas Experience

Michael Hicks
Shenandoah University

Mckenzie Boyce
Oklahoma State University

Brian Whitacre
Oklahoma State University

Dagney Faulk
Ball State University

Received: 04/24/2026
Accepted: 10/30/2026

Abstract

This study provides a causal analysis of the impact of data center expansion on local labor markets in Texas from 2000 through 2024. Using a panel of 254 counties, we employ staggered Difference-in-Differences (DiD) and event study frameworks, supplemented by Callaway and Sant'Anna (2021) diagnostics, to estimate the effects of these capital investments on employment growth. While traditional input-output (I-O) models often project significant job multipliers, our causal estimates find that data centers experience negligible long-term increases in labor demand. We find no statistically significant impacts to employment growth in Information (NAICS 51) or Professional and Technical Services (NAICS 54) sectors beyond a transient, pre-operational pulse and only a short-term increase in Construction (NAICS 23) employment growth following treatment. Furthermore, neither large nor small data centers appear to act as anchors for high-tech employment clusters, likely due to the proximity-agnostic nature of AI training workloads. Given that Texas tax incentives for these facilities are projected to cost $9 billion by 2030, our findings suggest that such fiscal outlays are an inefficient mechanism for local job creation.

1 Introduction

The rapid expansion of digital infrastructure, primarily data centers, has altered the focus of local capital attraction efforts in the United States. Although data centers are under construction or planned across most of the country, Texas has emerged as a leading destination for data center investment, driven by expansive land availability, energy resources, and fiscal incentives (Lin et al., 2024). As these facilities proliferate, their impact on local labor markets has become a subject of increasing policy concern.

State and local leaders often champion data center projects as catalysts for increased tax bases and jobs, including high-tech employment. Yet despite the scale of these investments, the empirical reality of their long-term labor market returns remains ambiguous. While construction phases are likely to generate temporary employment or bring in out-of-area workers that may generate additional local sales tax revenue, the extent to which data centers produce sustained job growth after becoming operational is less clear. This

study addresses this gap by providing a causal analysis of the employment effects associated with data center arrivals across Texas counties.

Understanding these impacts requires considering two competing approaches. From the perspective of agglomeration economies, the entry of data centers may signal a region's technological maturity. Such signals could potentially attract specialized human capital and generate knowledge spillovers in related sectors such as telecommunications and professional services. This perspective positions data centers as an anchor for a broader ecosystem of high-technology activity. A recent Pew Research survey suggests that many Americans agree with this view: 34% of respondents who had heard about data centers said that they were "mostly good" for local jobs, compared to 20% who said they were "mostly bad" (Gramlich et al., 2026).

In contrast, data centers are highly capital-intensive and require little labor during operations. Once the initial construction phase is completed, the operational phase typically requires minimal permanent staff. Furthermore, the substantial energy and land requirements of these facilities may impose localized costs. Strains on public infrastructure or upward pressure on utility prices may reduce the competitiveness of more labor-intensive manufacturing or industrial sectors, producing "crowding out" effects that offset any direct employment gains.

These competing dynamics are further complicated by the evolving geography of artificial intelligence (AI). A McKinsey & Company (2025b) analysis projected that by 2030, only about 30 percent of AI use will require physical proximity (within 20 miles) to a data center. This proximity-dependent share is comprised of real time inference, rather than training or storage of data. Thus, the regional economic effects of AI infrastructure are increasingly predicated on the distinction between training or storage and inference workloads.

Texas has successfully captured a large share of the global AI training market due to competitive energy pricing, existing infrastructure, and policy (Kenney, 2025). However, these facilities are often geographically divorced from the high-skilled labor markets required for AI inference that agglomeration economies would influence (Lin et al., 2024). This raises a central question for regional development policy: does the proximity-agnostic nature of modern AI training undermine the localized employment effects associated with high-technology investments? This study investigates that question by examining how data center investments affect labor market outcomes across Texas counties.

A key challenge to estimating these effects is the endogenous nature of firm location decisions. Data center operators systematically select jurisdictions that possess specific pre-existing advantages, such as proximity to fiber-optic trunk lines or favorable political climates for tax abatements (Kheybari et al., 2020). Failure to account for this selection process risks conflating pre-treatment advantages with post-investment outcomes. To overcome this selection bias, this research employs a staggered Difference-in-Differences (DiD) estimator (detailed in Section 3.2) to compare labor market trajectories between Texas counties that host data centers and a set of similar control counties without these investments. Focusing the analysis on Texas ensures the treated and control groups face the same state tax environment. The DiD framework is supplemented with an event-study design which allows for the detection of potential pre-trends and the decomposition of the treatment effect into distinct construction and operational periods, thereby providing a more dynamic understanding of the durability of employment gains.

This study contributes to the literature on regional infrastructure and industrial policy in several ways. First, it offers an updated empirical baseline for the Texas data center market, the second largest in the U.S.[1], using data through 2024 to capture the most recent wave of AI-driven expansion. Second, by focusing on sectoral employment shifts, the analysis identifies whether observed employment growth represents a net increase in regional opportunity or a mere reallocation of labor from traditional industries. In addition to filling a critical gap in the academic literature, these contributions are especially timely for policymakers and communities that need to evaluate several economic considerations of data centers (Boyce and Whitacre, 2026).

The remainder of this paper is structured as follows: Section 2 provides a review of the relevant literature on data centers and regional growth; Section 3 details the data sources and the econometric identification strategy; Section 4 presents the empirical results and robustness checks; and Section 5 offers a discussion of policy implications and directions for future research.

[1] As of 2025, Texas ranked behind only Virginia in terms of number of operational data centers (Batson, 2025).

2 Literature Review

The pace of investment in data center infrastructure indicates broad expectations regarding future demand for artificial intelligence (AI) and cloud-based services. Industry projections suggest that utility load may increase fivefold over the next two years, with data centers comprising the single largest component of growth in national load forecasts (Gamache et al., 2025). Projected data center growth rates through 2030 reflect a significant range of uncertainty, with estimates spanning from 4 percent to 166 percent (Makholm et al., 2024).

Given the surge in data center siting and expansion, academic research lags in being able to help nearby residents, policymakers and communities make informed decisions. However, there is a growing body of literature examining the localized effects of data centers. For example, preliminary work in Northern Virginia reported increased noise levels, escalating energy prices, and mixed fiscal outcomes associated with data center entry (Ngata et al., 2025).

Economic work has predominantly considered the effect of data centers on housing prices (Chia et al., 2026; Priest, 2026; Rubinovitz, 2026; Waters and Clower, 2025), with one additional working paper considering other economic effects such as local employment, wages, local income, and business establishments (Yue and Zeng, 2026). There is also a growing body of work focused on energy and environmental effects of data centers (for example, see Feher et al. (2025); Knittel et al. (2025); Kollar and Grady (2025); Siddick et al. (2021)). The effects of data centers on these various local factors remain largely mixed, further contributing to the uncertainty surrounding data centers and pointing to the need for more rigorous research.

2.1 Employment Effects

Literature on labor market effects of data centers is both new and sparse. Three key working papers have begun strengthening the academic literature: Alvarez et al. (2026), Bahar and Wright (2026), and Tian et al. (2026). Using a sample of US counties from 1995 to 2020, Alvarez et al. (2026) identifies the entrance decision of the data center operator using a Bartik-type shift-share instrument. They report enduring positive shocks to local economies, with total employment rising by 3.9% in counties with data centers. They test for effects beyond the technology and construction sectors and find overall employment growth is roughly unchanged when omitting the sectors hiring data center workers. They test the robustness of these findings with tax revenue and earnings data, reporting similar magnitudes of growth.

Bahar and Wright (2026) use a synthetic control method to identify the employment effect of data center openings. They report employment gains over a roughly 6-year period of up to 5%, with unequal effects in tech employment by data center type. Tenant data centers show no direct employment growth of technology jobs but retain overall employment gains over a long observation period. They also report wage growth because of data center openings. Tian et al. (2026) perform a synthetic control study of a single data center cluster in Oregon. They report large increases in construction employment during the construction phase, but no net employment gains after this phase.

Beyond these recent working papers, analysis has largely been limited to descriptive reports or input-output (I-O) analyses, often sponsored by data center developers or interest groups (Diao and Parker, 2025; Ferreira et al., 2026; Morrow County, 2024; PricewaterhouseCoopers, 2023b; Sage Policy Group, 2024). The I-O methodology typically estimates gross rather than net effects, which is equivalent to assuming perfectly elastic labor markets where every new data center job results in a net gain for the region (Thompson, 2019). This restrictive assumption limits the utility of these models in evaluating policies intended to influence business location (U.S. Chamber of Commerce Technology Engagement Center, 2017). Table 1 summarizes the findings of selected regional and national studies, incorporating standardized impacts for construction and operational phases.

Table 1: Input-Output Models of Data Center Effects

Study	Location	Direct Construction Jobs	Direct Operational Jobs	Total Employment (incl. multipliers)	Notes on Methodology / Multipliers
ECONorthwest (2014)	Prineville Data Center, Oregon	N/A	N/A	3,592	Historical ECONorthwest model evaluating infrastructure growth in rural counties
U.S. Chamber of Commerce Technology Engagement Center (2017)	"Typical" Facility	1,688 (local phase)	157 (annual)	N/A	RIMS II; typical 165k sf facility
Thompson (2019)	Sarpy County, NE	9,627 (Omaha MSA)	430	1,899 (4.4x statewide)	IMPLAN; reflects wind energy use
Mangum Economics (2023)	Chicago Area	8,000 (direct)	N/A	N/A	Employment effects as part of the fiscal footprint of the 2019 Illinois data center tax incentive.
PricewaterhouseCoopers (2023b)	USA National	1,000–10,000 per campus	50–300 per campus	6 × direct jobs	Descriptive with multipliers
Joint Legislative Audit and Review Commission (2024)	Virginia	3,000–5,000	100–400	6,000–8,000	Includes indirect/induced jobs
Sage Policy Group (2024)	Frederick Co, MD	∼2,000	∼200	3,000–4,000	IMPLAN multipliers
Morrow County (2024)	Morrow County, OR	1,500	80–120	∼2,500	Scenario modeling; not causal
Justice and Shepherd (2025)	Georgia	28,350	5,471	N/A	input-output modeling assessing the effect of $3.4B in construction output
SRC Evalmetrics (2025)	Ohio	19,537	17,320	95,000	All types of employment
Diao and Parker (2025)	Central and Southeastern Florida	21,449	2,487	34,025	Includes adjustment for tight labor markets (job cannibalization)
Ryu and Hiatt (2025)	US Forecast (2025-2030)	60,000 at peak (in 2026)	110,000 (in 2023)	170,000	Assumes levels of construction and operation employment by size.
Ferreira et al. (2026)	Great Lakes Region	33,459	4,356	78,970	S&P data & IMPLAN; this is cumulative for all facilities in the region connected to the grid
Roychoudhury and Ross (2026)	Northwest Indiana	5,393; 10,787; 16,180	463; 926; 1,388	2.26 direct and indirect; 3.22 including induced	3 hypothetical hyperscale configurations: 0.5, 1 and 1.5GW (5,10 and $15 billion in capital investment; RIMS II)
Mangum Economics (2026)	Virginia	N/A	N/A	112,000	IMPLAN model assessing the labor market effects of data center investment

Note: Studies highlighted in this table reflect the known input-output analyses on data center effects at the time of this publication. Given the increased attention on data centers and resulting proliferation of studies, this should be viewed as a select example of early input-output studies on data centers rather than a comprehensive inventory.

The employment estimates from these studies span a large range, suggesting a great deal of heterogeneity in the labor market effects of data centers. Some of these studies examine a single data center, with actual — or at least claimed — job creation data. Other studies provide national or statewide estimates. From a strict policy context, the gross job flows are probably not important. However, it seems likely that if these gross job flows are correct, then compensation or employment amenities of data centers represent an improvement over competing jobs within these locations. Still, the inherent weakness of I-O models for evaluating the likely employment effects is of concern in these estimates.

One way to judge the efficacy of these estimates is simply to evaluate the claims of one of these studies. The PricewaterhouseCoopers (2023a) study claims 50 to 300 direct operational jobs per data center. There are currently between 3,400 and 5,700 data centers in the U.S. Using the low- and high-end ranges of these estimates reveal 170,000 to 1.7 million direct jobs associated with data centers. From 2010 to 2024, jobs in NAICS 518 (Data Processing, Hosting and Related Services) grew by 256,000, but the parent sector NAICS 51 (Information) saw growth of only 160,000 jobs, suggesting many data center jobs might simply be due to shifting NAICS definitions within the Information industry, mirroring the well documented shift of manufacturing NAICS employment during the China Shock period (Bloom et al., 2019). Altogether, this implies that data center associated employment change is ephemeral and that net employment is near or below the lower bound of most of these studies. This provides further evidence of the need for a more rigorous analysis of the employment effects of data centers. The wide range in employment effects from these I-O analyses leave too much ambiguity for communities and policymakers making decisions about data centers and questioning the extent to which data center approvals might result in additional jobs in their area.

2.2 Economic Geography

Recent research provides a theoretical framework for understanding the location strategies of data centers. Arzumanyan et al. (2025) conduct a suitability analysis for data center placement in Texas considering local energy access, skilled labor, infrastructure, natural hazards, and land availability. They find that urban areas are generally more suitable for data centers, with energy access, infrastructure, and natural hazards being the most influential factors that impact site selection. It is notable that while skilled labor is considered, it does not emerge as the most influential factor in site suitability, perhaps foreshadowing a minimal effect of data centers on local employment.

Other work also distinguishes between third-party providers and hyperscale cloud providers. Third-party firms demonstrate a pervasive urban bias, locating near information-intensive sectors like finance, insurance, and real estate to reduce latency and satisfy consumer distaste for distance (Fang and Greenstein, 2025).

In contrast, hyperscale and footloose providers are far more sensitive to supply-side input costs, particularly industrial electricity prices and construction wages (Fang and Greenstein, 2025) or fiscal incentives (Gargano and Giacoletti, 2025). This allows hyperscale firms to concentrate their facilities in a small number of optimized, low-cost locations while serving national or international markets (Leka, 2022).

One implication of this is that smaller-scale facilities located near the appropriate cluster of human capital might represent the technology-based employment nodes. This would help explain the clustering of data centers in Loudoun County, Virginia, where land prices and property taxes are among the highest in the nation, but 72 percent of adults possess a bachelor's degree and 26 percent have a graduate degree. This also explains why Gargano and Giacoletti (2025) report that small data centers are largely unaffected by incentives.

2.3 Fiscal Incentives

A critical area of regional policy debate involves the fiscal incentives offered to attract data centers. State-level sales and use tax exemptions have led to significant revenue losses that are increasingly difficult for states to control (LeRoy and Tarczynska, 2025). Texas serves as a primary example of this trend. In early 2023, the state Comptroller projected data center subsidies for fiscal year 2025 would cost $130 million. By January 2025, this projection was revised upward to $1 billion. The Comptroller now estimates the total cost of the program will reach $9 billion between 2025 and 2030 (LeRoy and Tarczynska, 2025). Beyond sales tax exemptions, some states and communities compete for data centers by offering property tax abatements (Vaughan, 2025; National Conference of State Legislatures, 2025). This can impact school district and county / city revenue, which are heavily dependent on property taxes.

Gargano and Giacoletti (2025) analyze the effectiveness of U.S. state tax incentives in attracting data center investments and their subsequent economic impact. The authors find that incentives significantly increase data center construction, especially among large-scale facilities for cloud computing and large language models. These authors report that subsidies primarily alter the geographic distribution of infrastructure by increasing the project's Net Present Value, suggesting limited long-term spillover benefits for the local workforce.

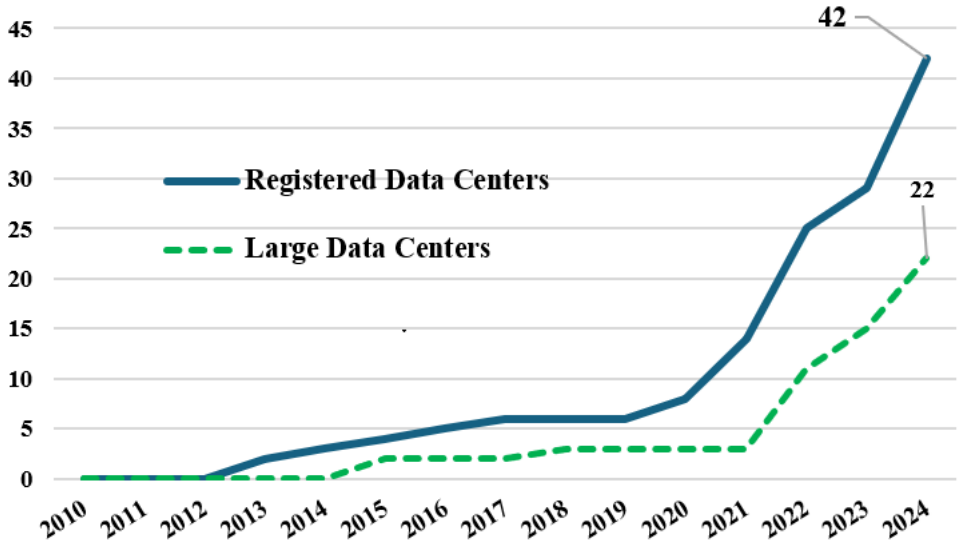
Skepticism regarding the necessity of these incentives is growing among researchers and taxpayers. Industry executives have noted that site selection is primarily driven by power availability, fiber connectivity, and infrastructure rather than tax breaks, with one Microsoft executive stating in 2024 that they could not recall a single placement decision determined by incentives (Nguyen and Green, 2025). It is fair to question this statement given the high-powered empirical results of Gargano and Giacoletti (2025). Despite this uncertainty, states continue to compete through fiscal incentives, with Virginia reporting nearly $929 million in state and local revenue losses for fiscal year 2023 (LeRoy and Tarczynska, 2025). These findings suggest that the industry provides significant capital investment, but the long-term fiscal and labor-market impacts require more rigorous, causal evaluation beyond traditional multiplier-based frameworks.

3 Data and Empirical Methods

3.1 Data Sources and Variable Construction

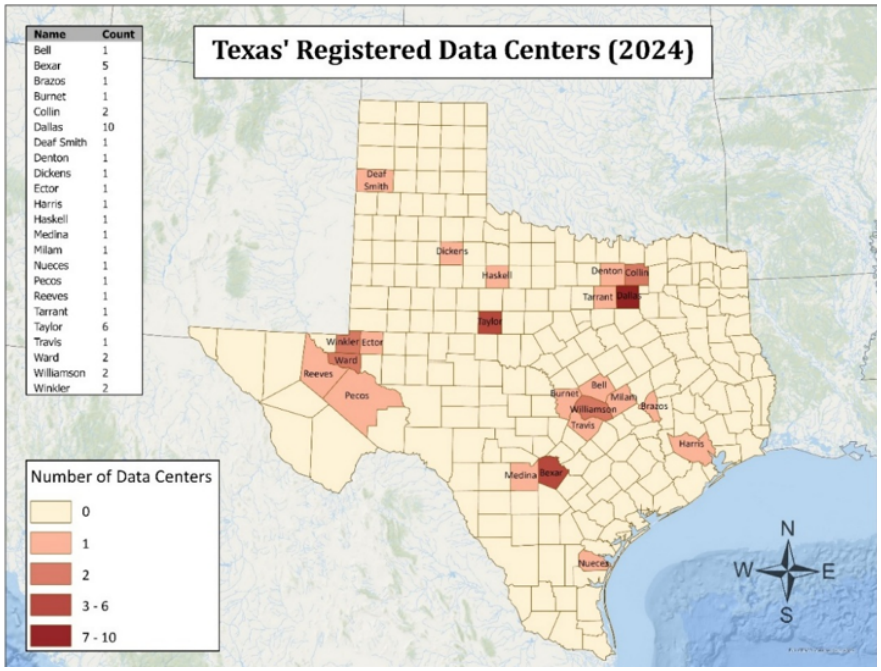
This study utilizes a panel dataset of 254 Texas counties spanning 2000Q1 through 2024Q4, yielding 25,400 potential observations, with 64 total data centers in our sample. Texas serves as an ideal laboratory for this analysis due to the high variance in data center adoption by counties and the transparency of its regulatory environment. Treatment data are sourced from the Texas Comptroller of Public Accounts, which maintains granular records of "Registered" and "Large" data centers which have applied for and received state tax incentives. Registered data centers (42 in our sample) are classified as less than 100,000 square feet and large data centers (22 in our sample) are defined as those exceeding 100,000 square feet. This size differentiation is of interest because registered and large data centers likely differ in their capital intensity, energy demand, land use, and economic impacts. In Figure 1, we illustrate the cumulative number of data centers in Texas from 2010-2024. We map the registered and large data centers as of 2024 by county in Figures 2 and 3. Note that while some overlap between registered and large data center locations exist (for example, there are clusters of both in Taylor County), there are many counties with only one type.

Figure 1: Cumulative Data Centers in Texas, 2010-2024



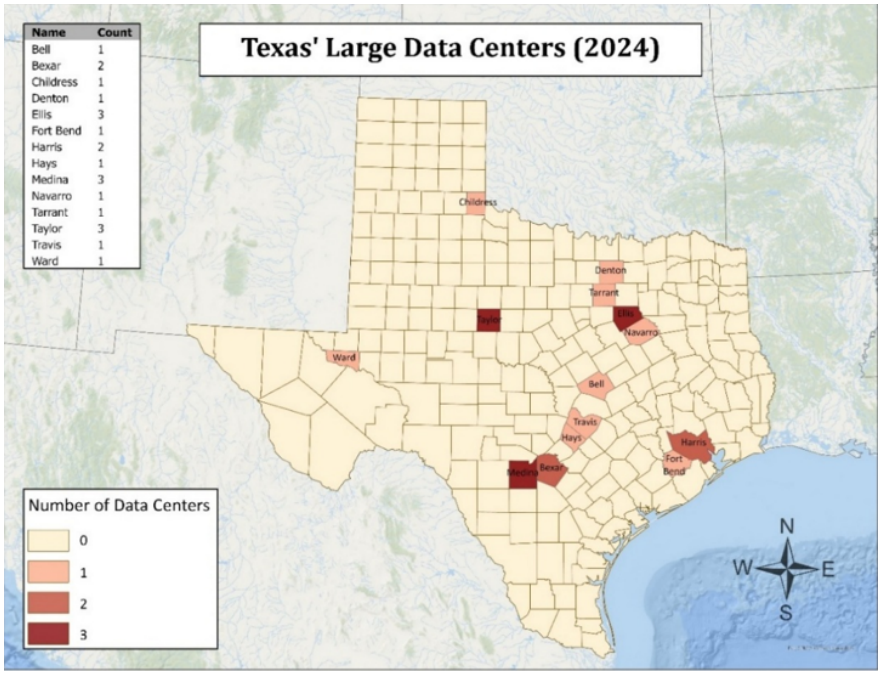
Note: Registered data centers are those less than 100,000 square feet and large data centers exceed 100,000 square feet. Source: Texas Comptroller of Public Accounts

Figure 2: Registered Data Centers in Texas (2024)



Note: Registered data centers are less than 100,000 square feet and large data centers exceed 100,000 square feet. Source: Texas Comptroller of Public Accounts

Figure 3: Large Data Centers in Texas (2024)



Note: Registered data centers are less than 100,000 square feet and large data centers exceed 100,000 square feet. Source: Texas Comptroller of Public Accounts

Our treatment variable is constructed using the effective dates of Texas data center personal property tax exemptions. Specifically, we identify the county and effective date for each certified data center receiving

a tax exemption under Texas Tax Code Section 151.359 or 151.3959.[2] We define treatment timing as the quarter / year in which a data center's tax exemption becomes effective. Throughout the analysis, this date should be interpreted as the commencement of certified investment activity rather than the date a facility becomes operational.

We use the effective date because it is the only project milestone that is observed consistently and precisely for all qualifying data centers in Texas. Alternative milestones, such as project announcement dates, construction groundbreaking dates, and operational dates, are not systematically available across facilities. The effective date of the data center tax exemption is also reflective of the investment activity that generates local labor demand. Under Texas law, data centers can apply the exemption immediately upon certification by the Comptroller, incentivizing operators to receive the exemption prior to any major capital investment. However, this certification also becomes public record making it likely to either coincide with or fall after public announcement of the data center. Additionally, firms must satisfy specified employment and investment thresholds within five years of the effective date, and the exemption expires after 10-20 years, both of which incentivize operators to apply for the exemption immediately prior to employment and capital investments to maximize the benefit of the exemption. While some preliminary work, such as site selection and permitting, may take place prior to this effective date, we anticipate that would be done under non-disclosure agreements and predominantly with non-local employees and would not clearly create an anticipatory effect on local employment. The construction duration will vary by data center size, with industry publications suggesting data center construction typically takes between six and 36 months.[3] The long construction period could create a lagged effect but still falls primarily within our observation window.

Labor market data are derived from the Quarterly Workforce Indicators (QWI). We examine employment effects across several North American Industry Classification System (NAICS) hierarchies to capture both direct and indirect effects. Specifically, we analyze: Information Sector (NAICS 51), Construction Sector (NAICS 23), Professional Services (NAICS 54) along with total employment.[4] These sectors were selected given the notion that data centers will bring vast construction employment during the building phase and high-tech jobs in the operational phase (McKinsey & Company, 2025a). Summary statistics for counties with and without a data center are provided in Table 2. Summary statistics for all counties are included in Appendix Table A1.

Table 2: Summary Statistics (2000-2024)

	Counties with a Data Center			Counties without a Data Center		
Employment – Levels	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Total	381,821.6	65,099	534,334.1	30,668.7	3,736	149,175.6
Construction	22,863.5	5,079	31,236.4	2,260.3	261	11,182.8
Information	9,685.1	1,023	15,634.3	797.1	36	4,409.9
Professional & Technical	35,024.9	2,606	57,673.5	2,161.9	100	14,018.9
Employment – Growth Rate						
Total	0.6%	0.7%	3.2%	0.4%	0.4%	6.4%
Construction	0.8%	0.7%	16.3%	0.7%	0.6%	19.8%
Information	0.1%	0.0%	27.2%	-0.02%	0.0%	29.9%
Professional & Technical	1.2%	0.9%	19.2%	0.9%	0.6%	20.1%

These data illustrate large differences in employment levels between the average county with and without data centers. This extends across all types of reported employment. Employment growth is more similar,

[2]This limits the sample to data centers meeting, or anticipating to meet, minimum employment thresholds (20 full-time, permanent employees making 120% of the county's average weekly wage for registered data centers and 40 for large data centers) and investment thresholds ($200 million for registered data centers to receive a 10-year exemption, $250 million for registered data centers to receive a 15-year exemption, and $500 million for large data centers to receive a 20-year exemption).

[3]See https://broadstaffglobal.com/data-center-construction-timeline and https://www.theguardian.com/technology/2026/jul/07/stymied-datacentre-projects-threaten-global-ai-revolution.

[4]Of the 25,400 potential observations noted above, some county-quarter outcome variable observations are missing due to QWI data suppression. The missing observations are 13 (0.051%) for total employment, 393 (1.54%) for Construction Employment, 2,205 (8.68%) for Information Employment and 794 (3.13%) for Professional & Technical Services Employment. The resulting panels are not necessarily balanced.

both in levels and in variation (standard deviation). The one notable example is the information sector, which grows slowly in counties that hosted a data center, and is in decline elsewhere.

A number of site-selection factors identified in the data center location literature, including industrial electricity prices, water availability, and proximity to fiber-optic infrastructure, are not consistently observable at the county level over our 2000–2024 sample period. Industrial electricity rates in particular are typically negotiated at the establishment level (Byrne et al., 2022) and are not systematically reported to researchers. We are therefore unable to include these factors directly as covariates. Section 3.2 discusses how our identification strategy addresses this limitation.

3.2 Econometric Identification

Given that data centers were established at various points throughout the sample period, we adopt a generalized staggered Difference-in-Differences (DiD) specification which provides a causal estimate of the labor market impact of data center deployment:

$$dln(N_{s,i,t}) = \alpha + \beta D_{i,t} + \gamma_i + \gamma_t + \epsilon_{i,t}, \tag{1}$$

where $dln(N_{s,i,t})$ is the difference in the natural log of employment from the previous quarter — i.e., the employment growth rate — in sector s, county i, and quarter t. We also test for employment levels $ln(N_{s,i,t})$ as a robustness check, with results provided in Appendix Tables A3 – A6. We add 1 to $(N_{s,i,t})$ to include counties with zero employment in specific sectors. The variable $D_{i,t}$ is a count treatment variable representing the cumulative number of data centers in a county that have applied for and received a tax exemption from the Comptroller. This is identical to the count specification in both Alvarez et al. (2026) and Bahar and Wright (2026). Unlike a binary indicator, the count specification captures the marginal intensity of additional infrastructure investments. This allows the estimated effect to reflect incremental changes in employment associated with each new data center investment, rather than a single average treated-untreated difference. The parameters γ_i and γ_t represent county and quarter-year fixed effects, respectively, which control for time-invariant county characteristics and state-wide macroeconomic shocks. This ensures identification comes from within-county variation in data center intensity over time, rather than from persistent cross-county differences or common economic trends.

This design directly addresses the selection concern raised in Section 1: that data center operators may favor counties with unobserved locational advantages, such as proximity to fiber-optic trunk lines or existing water and power infrastructure. To the extent that these advantages are stable over our sample period, county fixed effects (γ_i) absorb them: a county's baseline endowment of fiber infrastructure or water resources is unlikely to change meaningfully within a 24-year window, so it does not bias the treatment coefficient. The more difficult case is a time-varying siting advantage — for example, a county-specific shift in industrial electricity prices coinciding with data center entry. Because such data are not systematically observable at the county level (see Section 3.1), we cannot include them as covariates or implement a doubly robust design incorporating time-varying propensity scores. Our identification therefore rests on the parallel trends assumption rather than on selection-on-observables: we require that, absent treatment, treated and control counties would have followed similar employment trajectories, not that they be identical in levels.[5] We assess this assumption directly via the pre-treatment leads reported in Section 4.2 and the parallel-trends test statistics reported alongside each specification in Section 4.1.

The coefficient of interest, β, identifies the average treatment effect on the treated (ATT) under the parallel trends assumption: that in the absence of a data center arrival, the employment trajectory of treated counties would have evolved identically to that of the control group.[6] We make an additional assumption known as the stable unit treatment value assumption (SUTVA). This assumes county i's employment is unaffected by other counties' receiving a data center. Thus, in the context of our modeling, this assumption precludes the possibility of spatial spillovers of data center employment, so that the construction of a data

[5] Note that the parallel trends assumption tests for trajectories in growth rates when our dependent variable is $dln(N_{s,i,t})$ and for trajectories in employment levels when our dependent variable is $ln(N_{s,i,t})$

[6] See Hicks and Faulk (2024) for an example of a place-based Difference-in-Differences estimate using this specification.

center in county i does not affect employment growth in county j. Delgado and Florax (2015) recommend a spatial model to account for both direct and indirect treatment effects.

We expand the Delgado and Florax approach, as in Hicks et al. (2026), which accounts for the spillovers of data centers on adjacent counties. The model takes the form:

$$dln(\hat{W}N_{s,i,t}) = \alpha + \beta D_{i,t} + \gamma_i + \gamma_t + \epsilon_{i,t}, \tag{2}$$

where $\hat{W}$ is a first order queen contiguity weight matrix, excluding non-Texas border counties, and $D_{i,t}$ is the treatment variable indicating the count of data centers in county i and period t. In this case, rejecting the null on β rejects the assumption of SUTVA. In testing across the full set of dependent variables, we find no statistically significant evidence of spillover effects of a data center on neighboring employment. These results are reported in Appendix Table A2.

3.3 Event Study Design and Dynamic Effects

Recognizing the potential for heterogeneous treatment effects over time, particularly the distinction between short-term construction booms and long-term operational plateaus, we supplement the static DiD with an event study specification (Callaway and Sant'Anna, 2021). This approach proceeds in two stages. It first estimates group/time average treatment effects by cohort (simultaneous treatment). This compares never treated and not yet treated observations to avoid comparing a newly treated county (those with a new data center) to a county which had previously had a data center. The second stage aggregates these cohort effects into a single dynamic event study avoiding comparing new treated observations against previously treated observations.

This allows us to test for pre-treatment trends and map the dynamic labor market response:

$$dln(N_{s,i,t}) = \lambda + \sum_{k=-j}^{m} \beta_k D_{i,t-k} + \gamma_i + \gamma_t + \epsilon_{i,t}. \tag{3}$$

In this model, the coefficients β_k represent the leads and lags relative to the quarter of treatment ($t = 0$). The leads ($k < 0$) test for statistically significant coefficients prior to treatment, which might indicate a violation of the parallel trends assumption (i.e., that data centers are selecting into counties already experiencing idiosyncratic employment growth). The lags ($k > 0$) reveal the durability of the employment effect, allowing us to observe, for example, if initial labor gains in the construction sector persist or mean-revert following facility completion.

4 Estimation Results

4.1 Difference-in-Differences Estimates of Data Centers

In Table 3, we report the results of the data center count difference-in-difference estimates. These were estimated by running Equation 1 on four dependent variables: growth rates in (1) total employment, (2) Construction employment (NAICS 23), (3) Information Services employment (NAICS 51), and (4) Professional and Technical Services employment (NAICS 54).

We also report the parallel trends assumption test statistic, and in each case, we were unable to reject the assumption of parallel trends between the control and treatment groups. The DiD specification in this data center entrance context is preferred over the Bartik-style Instrument of Alvarez et al. (2026) because it requires fewer, and more readily tested, assumptions about the initial equality of the treatment and control groups. The IV approach requires exogeneity from the local and global shock shares. This is difficult to directly test (Goldsmith-Pinkham et al., 2020). The synthetic control method likewise requires a weighted combination of control observations, rather than simple parallel trends that are sufficient and easily tested in a DiD framework. Both methods have their place, particularly regarding uncertainty about facility opening

dates (as distinct from our tax-exemption-based treatment measure) that plague much of the commercially available data, and all the publicly available data outside of Texas.

Table 3: Employment Growth Rate Difference-in-Differences Results (2000–2024)

	Total Employment	Construction Employment	Information Services Employment	Prof. & Tech. Services Employment
Data Centers (β)	0.001**	-0.002	-0.001	-0.0001
	(0.0004)	(0.001)	(0.003)	(0.001)
County FE (i)	Yes	Yes	Yes	Yes
Quarter FE (t)	Yes	Yes	Yes	Yes
N	25,130	24,714	22,873	24,323
R^2	0.06	0.02	0.01	0.02
Parallel Trend Statistic	0.34	0.15	0.16	0.43

Note: Treatment is a count of all data centers in county i, quarter t. Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023). * and ** denote statistical significance at the $p < 0.10$ and $p < 0.05$ levels, respectively.

The empirical results presented in Table 3 provide an estimate of the impact of data center expansion on employment growth in Texas counties. The most striking finding is the statistically significant, positive coefficient for total employment. We find each new data center adds only 0.08 percentage points to county employment growth rates on average across the post-treatment period. This is too small to meet a minimum threshold of economic significance, and as the duration results in Table 5 will show, the effect does not result in a shift in the total employment level over time. Graphical depiction of these DiD estimates for average treated vs. control growth rates are in the Appendix as Figures A1-A4. We discuss this result more as we consider the staggered nature of the treatment.

The sectoral decomposition reveals that there is no statistically discernible demand for labor within Construction, Information Services, or Professional and Technical Services due to these data centers. This challenges an interpretation of vast construction and high-tech employment growth associated with data centers. The estimate for construction is negative and insignificant, suggesting that the initial build-out of these facilities does not result in a sustained or even detectable increase in local construction employment growth rates over the longer term. This likely points to a cannibalizing of existing construction workers from residential or commercial projects or the use of specialized, non-local contracting firms or a "pulse" of activity too brief to shift quarterly county-level data. This could occur from temporary, transient specialized construction workers – which will almost certainly be the case when the number of construction jobs exceeds the capability of the local workforce (for example, the thousands of direct construction jobs noted in Table 1 taking place in a remote county).

More importantly, the near-zero coefficients for Information and Professional Services suggest that the presence of data center infrastructure fails to catalyze a local technology cluster. Despite the massive scale of these facilities, they do not appear to act as anchors for growth in the high-skilled service jobs typically associated with the digital economy. This aligns with the argument that the physical infrastructure is highly automated and geographically decoupled from the technical talent that manages the software and systems residing within the data center.

From a methodological standpoint, the parallel trend statistics offer some support for the study's internal validity. Across all four models, the p-values for the parallel trend tests are well above 0.05, ranging from 0.66 to 0.87, and we fail to reject the null hypothesis of common pre-trends in every instance. This is consistent with comparable pre-treatment trajectories between treated and control counties. We note, however, that a failure to reject the null reflects the power of the test as well as the true absence of differential trends, and so should be interpreted as reassuring rather than confirmation of the identifying assumption. Taken together with the sectoral results above, the estimates suggest that while data centers provide a small lift to total county employment growth, they do not currently serve as direct engines for localized growth in specialized information or technical service sectors.

To ensure the robustness of our staggered DiD estimates against potential treatment effect heterogeneity and the "negative weighting" problem associated with Two-Way Fixed Effects (TWFE), we employ the Callaway and Sant'Anna (2021) estimator. This allows for the identification of group-time average treat-

ment effects using only never-treated or not-yet-treated units as valid counterfactuals. The results for this robustness check are provided in Tables 4 (group/cohort effects) and 5 (duration effects), and we discuss them here.[7]

Focusing primarily on the total employment growth effects, the Callaway and Sant'Anna (2021) results provide a critical refinement to the previous findings. While the standard staggered Difference-in-Differences suggested a small, positive total employment growth effect (Table 3), these group-time and duration-specific estimates reveal that the impact is inconsistent across time.

The group effects presented in Table 4 demonstrate that the average impact of a data center depends on the quarter of treatment. Since these estimators are comparisons of treated counties versus never treated and not yet treated counties, the lack of coefficient estimates in some year/quarters reflects an insufficient number of newly treated counties, in both not-yet treated and never treated groups. For cohorts that received facilities between 2012 and 2022, the coefficients are statistically insignificant, suggesting that for over a decade of the sample, data center entries had no detectable impact on county-level employment growth. In the 2023Q1 and 2024Q4 cohorts, there was a small negative effect. Since there is no meaningful post-treatment for 2024Q4 observation, this likely reflects one quarter of Construction Employment. Note that this is consistent with the negative sign for Construction in Table 3 and suggests that cohort-specific impacts can arise. It is possible that this volatility in more recent periods reflects large but localized disruption of hyperscale AI deployments, where the labor market experiences sharp shocks (potentially construction employment related) that vary widely by project scale. A more cautious interpretation is that the absence of any post 2024 observations offers little dynamic interpretation to these cohorts.

Table 4: Group Effects, Total Employment Growth Rate, Callaway and Sant'Anna (2021) estimators

Group	Coefficient	Std. Error
2015Q2	-0.005	0.007
2016Q2	-0.006	0.005
2020Q2	0.008	0.008
2020Q3	-0.001	0.013
2021Q2	-0.013	0.008
2021Q3	0.008	0.005
2022Q1	-0.005	0.005
2022Q2	-0.001	0.006
2022Q3	0.008	0.011
2023Q1	-0.010*	0.006
2023Q3	0.003	0.006
2024Q1	-0.004	0.003
2024Q3	0.015	0.003
2024Q4	-0.005***	0.007

Note: N/A observations lack sufficient observations for standard error calculations, due to lack of variation in treatment in those time periods. Standard errors are clustered at the county level. *, **, and *** denote statistical significance at the $p < 0.10$, 0.05, and 0.01 levels, respectively.

The duration effects shown in Table 5, which aggregate the average treatment effect on the treated counties by the number of quarters since treatment, offer perhaps the most powerful evidence against long-term growth associated with data centers. For the first 20 quarters (5 years) of treatment every coefficient is statistically insignificant. This implies that even the immediate construction pulse is not large enough to shift the quarterly trend in a statistically reliable way across the treated counties, suggesting that the initial promise of the facility never materialized into permanent local employment growth. It may also be that the observed employment effects at data centers simply relocated workers from other activities in the

[7]Group (or cohort) effects from the Callaway and Sant'Anna approach show the treatment effect separately for groups first treated in different calendar periods (e.g., the 2018Q4-treated cohort versus the 2020Q2-treated cohort). Duration effects instead aggregate across cohorts by time since treatment (event time), showing the average effect 1 year after treatment, 2 years after treatment, etc., regardless of when treatment began.

same sectors (construction, information and professional and technical services) without resulting in net employment growth.

Table 5: Duration Effects, Total Employment Growth, Callaway and Sant'Anna (2021) Estimators

Duration	Coefficient	Std. Error
0	-0.003	0.005
1	-0.003	0.006
2	0.001	0.006
3	0.000	0.006
4	-0.005	0.006
5	0.002	0.007
6	-0.006	0.007
7	-0.007	0.006
8	-0.006	0.006
9	-0.002	0.005
10	-0.001	0.006
11	-0.004	0.005
12	-0.002	0.006
13	0.001	0.008
14	0.002	0.007
15	0.001	0.008
16	-0.001	0.006
17	0.000	0.007
18	0.002	0.008
19	-0.002	0.004
20	-0.002	0.010

Note: N/A observations are pre-treatment periods. Standard errors are clustered at the county level.

When comparing these results to the earlier Two-Way Fixed Effects DiD model, the Callaway and Sant'Anna diagnostic clarifies why the previous total employment growth coefficient appeared positive. The previous result was likely influenced by comparisons where early-treated counties with declining trends were used as controls for later-treated ones.[8] Once the model uses clean controls consisting of never-treated or not-yet-treated units, the evidence for a positive employment effect largely vanishes. The labor market return on data centers in Texas appears to be noise rather than signal, characterized by transient shocks and long-term neutrality.

A Goodman-Bacon (2021) Decomposition[9] of our staggered DiD reveals that about 92 percent of the results are driven by the comparison of treated to never treated counties. This is not surprising given the large number of never treated cross sections in this sample.

To assess whether larger or smaller data centers might have different relationships with employment growth, we evaluate the effect of only large-scale data centers (greater than 100,000 square feet) and on smaller scale (Registered) data centers of less than 100,000 square feet on these four employment growth rate measures. The results from large data centers are shown in Table 6 and the results from registered data centers are shown in Table 7.

[8]This is referred to as forbidden comparisons by Callaway and Sant'Anna.

[9]A Goodman-Bacon Decomposition decomposes the TWFE estimate into a weighted average of all possible DiD comparisons to reveal which comparisons drive the estimate. This helps identify forbidden comparisons where already-treated units are used as controls, potentially introducing bias into the estimated treatment effects.

Table 6: Employment Growth Rate Difference-in-Differences Estimates from Large Data Centers in Texas (> 100,000 square feet), 2000–2024

	Total Employment	Construction Employment	Information Services Employment	Professional & Technical Services Employment
Coefficient (β)	0.001	-0.00003	-0.016*	-0.005
	(0.001)	(0.003)	(0.009)	(0.005)
County FE (i)	Yes	Yes	Yes	Yes
Quarter Time FE (t)	Yes	Yes	Yes	Yes
N	25,130	24,714	22,873	24,323
Adjusted R^2	0.05	0.001	-0.01	0.01
Parallel Trend Statistic	2.46**	0.11	0.63	0.33

Note: Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023). * and ** denote statistical significance at the $p < 0.10$ and $p < 0.05$ levels, respectively.

Table 7: Employment Growth Rate Difference-in-Differences Estimates from Registered Data Centers in Texas (< 100,000 square feet), 2000–2024

	Total Employment	Construction Employment	Information Services Employment	Professional & Technical Services Employment
Coefficient (β)	0.001**	-0.004	0.002	0.001
	(0.001)	(0.003)	(0.004)	(0.002)
County FE (i)	Yes	Yes	Yes	Yes
Quarter FE (t)	Yes	Yes	Yes	Yes
N	25,130	24,714	22,873	24,323
Adjusted R^2	0.05	0.01	-0.01	0.01
Parallel Trend Statistic	0.878	0.86	0.973	0.40

Note: Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023). * and ** denote statistical significance at the $p < 0.10$ and $p < 0.05$ levels, respectively.

This decomposition by data center size reveals a deviation from the aggregate effects. For large data centers, there is no statistically significant aggregate employment growth, Construction employment growth, or Professional and Technical services growth effect. We note that the parallel trends test for the Large Data Centers' effect on the total employment growth rate rejects the null at conventional levels (p = 0.014). This weakens confidence in a causal interpretation of that specific coefficient, and should be read with caution. The results do show a significant negative effect on Information Services, meaning an increase in large data centers in a county results in a decrease in the growth rate of Information Services employment. This could signal a "crowding out" effect, where labor-intensive local information service firms might be pushed out of the market through higher energy and infrastructure costs. Coupled with data centers requiring little within-county information service employment, this results in a net decrease in Information Services employment growth. An increase in small-scale (Registered) data centers results in a positive but non-meaningful increase in total county employment growth, mirroring the aggregate regression result in Table 3.

These tests of labor market effects due to the scale of the data center provide a similar finding to that of the aggregate data center estimates. In both sets of estimates, across both the employment growth and levels estimates the effects range from zero to minimal, with both low point estimates and statistical uncertainty.

As a robustness check, we also examined the potential clustering of large data centers by including a squared term for the data center count variable for the outcome measures reported in Table 6. In each case the magnitude of the coefficient dropped by roughly half or more, but none became statistically significant.[10] This offers no evidence of clustering, an issue we discuss later. These results are available upon request.

To further assess the granularity of the labor market response, we extend the analysis to highly disaggregated industrial classifications within the Information and Professional Services sectors. By shifting from broad 2-digit NAICS codes to more specific 3- and 4-digit subsectors, including Information (NAICS 517), Data Processing, Hosting, and Related Services (NAICS 518) and Computer Systems Design and Related Services (NAICS 5415) we attempt to isolate the precise locations where data center employment growth might manifest. Note that NAICS 517 is the subsector of the NAICS 51 Information sector focused specifically on telecommunications, and that NAICS 5182 Computing Infrastructure Providers, Data Processing,

[10] As an example, the squared data center count variable for growth in total employment growth yielded a coefficient of 0.00005, which is about 1/20th of the Table 6 coefficient.

Web Hosting, and Related Services, is nearly a perfect correlate of the broader NAICS 518 in Texas counties. The results from this disaggregated analysis using employment growth rate as the dependent variable are presented in Table 8.

Table 8: Disaggregated Employment Growth Rate Difference-in-Differences Estimates (2000–2024)

	Information (NAICS 517)	Data Processing, Hosting and Related Services (NAICS 518)	Computer Systems Design (NAICS 5415)
Coefficient (β)	-0.003	0.002	0.001
	(0.004)	(0.004)	(0.004)
County FE (i)	Yes	Yes	Yes
Quarter FE (t)	Yes	Yes	Yes
N	20,288	7,667	17,742
Adjusted R^2	-0.00	-0.01	-0.01
Parallel Trend Statistic	-1.28	-0.12	-0.82

Note: Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023).

In each instance, we find no evidence of a data center effect. The point estimates are small and statistical uncertainty is high. It is also worth noting that these estimates are plagued by omitted observations, with less than a third of counties appearing in the NAICS 518 category.

As recognized in Section 3.2, there is also likely interest in employment levels in addition to the impacts on growth rates reported above. Those results are shown in Appendix Tables A3 – A6. Generally, there are more problematic pre-trends when using the employment level data; however, we do document a marginally statistically significant 6.2% increase in Professional and Technical jobs and a more meaningful 16% increase in Construction jobs that shows up only for large data centers.

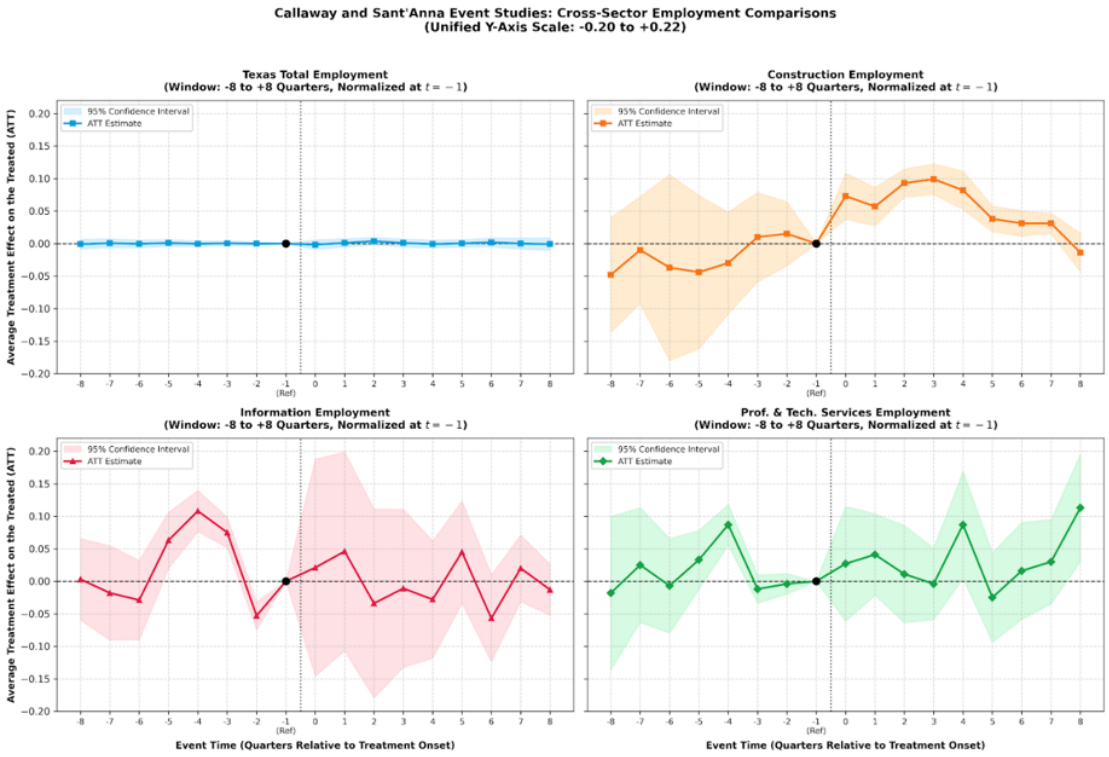
The Difference-in-Difference approach offers an average treatment effect of post-entrance of each new data center (or large data centers). These models do not capture the dynamic effects. For that, we turn our analysis to event studies.

4.2 Event Study Estimates of Data Centers

The Callaway and Sant'Anna (2021) difference-in-differences estimator can be extended to an event study, by disaggregating estimation into group-time average treatment effects, comparing specific timing cohorts to never-treated and/or not-yet-treated control groups at given time periods. We use both the treated and not yet treated counties as control groups.

By aggregating these group-time parameters into a unified event-study plot, the estimator yields dynamic treatment effects relative to the period of initial exposure without contamination from earlier cohorts. This event-study aggregation enables a clean pre-trend test, evaluating whether treatment and control units followed parallel paths prior to intervention, while also providing evidence of how treatment effects evolve, persist, or dissipate over relative time from treatment. Figure 4 provides these event-study plots for our four main employment sectors of interest.

Figure 4: Event Study Effects (Callaway and Sant'Anna, 2021) of staggered treatment: Growth Rates in Total Employment, Construction, Information and Professional and Technical Services Employment in Texas (2000:Q1-2024:Q2)



The visual comparison in Figure 4 reveals notable heterogeneity in treatment effect dynamics, magnitude, and precision across growth rates in total employment, Construction employment, Information employment, and Professional and Technical Services employment.

Prior to treatment onset, point estimates for total employment growth remain tightly clustered near zero with narrow confidence bands, supporting the underlying parallel trends assumption required for valid difference-in-differences identification. This is consistent with the static DiD parallel trend statistic provided in Table 3. Following treatment, total employment exhibits negligible movement away from zero, indicating little to no aggregate labor market expansion or contraction attributable to the intervention.

In contrast, sector-specific estimates demonstrate greater volatility and variation in magnitude. Construction employment growth shows a clear positive post-treatment trajectory, rising sharply in the initial quarters following treatment onset and maintaining elevated point estimates through t = 3 before declining in t = 4. This pattern is consistent with a transitory infrastructure or build-out effect, where physical development drives short-term labor demand increases for roughly 1 – 1.5 years before tapering off as projects reach completion.

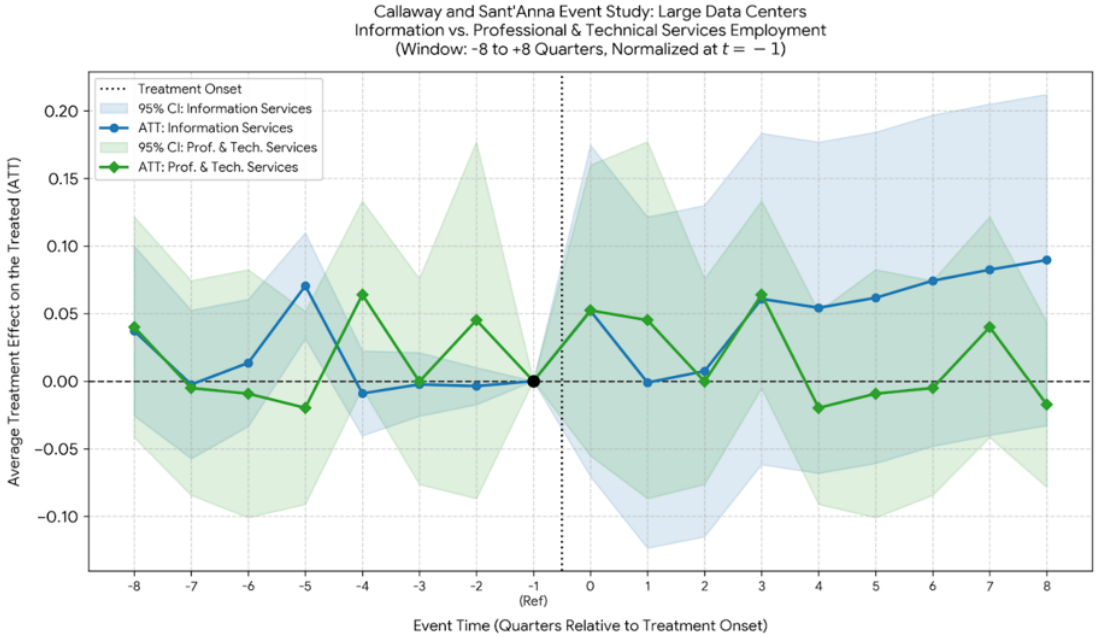
Conversely, the estimates for Information employment and Professional and Technical Services employment display substantial quarter-to-quarter fluctuation both before and after the event horizon. While Professional and Technical Services employment experiences periodic post-treatment spikes (note t = 4 and t = 8), the 95% confidence intervals expand post-treatment. This widening uncertainty likely reflects increased cohort-level effect heterogeneity over time. For the information sector, the coefficients are not significantly different from zero in any of the post time periods. For professional and technical services, the coefficient is positive and statistically different from zero only in the eighth quarter. As a result, while specific sector-level point estimates suggest localized labor adjustments, the wide confidence bands urge caution in inferring long-term, statistically precise structural shifts in the Information and Professional & Technical Services sectors.

We also conduct these event studies for employment in levels (Appendix Figure A.5). Across total employment, Information employment and Professional and Technical services employment, there were no visible cumulative employment changes in point estimates in the quarters following the data center entrance. As with the growth event study, Construction employment in levels showed some increases in the eight quarters after treatment. Although not statistically significant across any of these observations, the positive trend post-treatment reinforces the findings for the construction growth rate in Figure 4. However, the level of uncertainty in all employment levels point estimates rose substantially, reinforcing the evidence from the growth rate estimates that there is remarkable heterogeneity in the employment effects of data centers.

As with the growth estimates, it is possible there is another margin on which sectoral employment may be affected. For example, employment in accommodations, eating and drinking establishments, or real estate may be affected by the construction phase. Nevertheless, with these results in aggregate total employment no different from zero, it seems unlikely that more intra-county employment shifts will be observed in these data. As robustness tests, we further isolate our test of the clustering hypothesis of large-scale data centers by treating only large data centers, which might be associated with the low latency (distance) requirements discussed in the McKinsey & Company (2025b) study. We focus on the Information and Professional and Technical Services sectors.

Those results indicate that for at least a year after treatment, there is no evidence of increases in employment growth rates for Information Services, but point estimates after one year indicate a steady increase in employment growth, although not statistically different from zero. If we extend the post-treatment period to 16 quarters, we see only one observation ($t + 16$) that is different from zero. We note that there are only three observations of large data centers that experience 16 or more quarters post-treatment. For Professional & Technical Services, there is no significant evidence of employment growth rate effects through $t + 8$ or later, as shown in Figure 5. Insofar as Large Data Centers ($> 100,000$ square feet) are different from small data centers, this difference would likely manifest itself in labor markets as employment growth in either Information Services supporting data center operations, or in Professional & Technical services associated with AI or data center management. There is some weak evidence of an Information Services effect, but none for Professional & Technical Services.

Figure 5: Employment Growth Effects on Information and Professional Services of Large Data Center Count in Texas



Notes: 8 quarters pre/post filing, with 95% Confidence Interval.

The event study results continue to point to heterogeneity of post treatment results in all but Construction employment, where there is increased statistical certainty following the opening of a data center. The heterogeneity is most pronounced in the total employment estimates. However, in both the growth rate and levels estimates (reported in the Appendix), the variance inflation factors shows no evidence of collinearity, which might contribute to the higher variance post treatment. There are several possible explanations for this.

The least likely issue is unbalanced panels. Though unbalanced panels due to missing observations can contribute to variance inflation, the most heavily affected variable (total employment) has a tiny share of omitted observations (13 out of 25,400). Thus, we do not view this as a plausible issue.

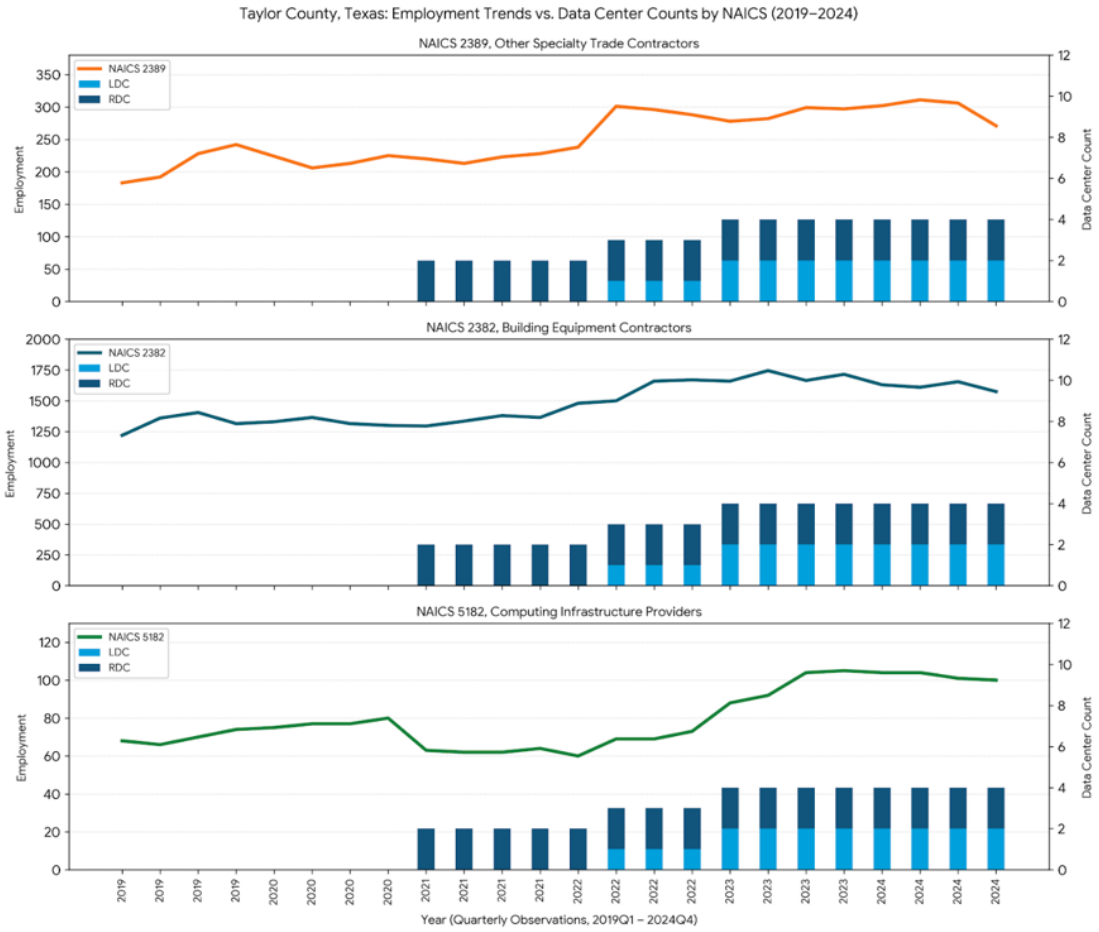
Second, our treatment dates may differ from actual construction start or opening dates. As we described above, we have a month and year in which the data center operator had a tax incentive application approved in Texas. Although the incentive is designed to promote close adherence to the date of approval, there is likely some variability in the actual start date of construction and operation. However, the fact that the variance in both total employment and Construction employment shows dramatic variability around the treatment date, we believe data error is not a contributing factor to these changes in variances. If anything, the changes in variance, in say Construction employment after treatment (and only a modest, one or two quarter change prior to treatment), suggest the data is substantively correct.

There may also be administrative limitations on the data which prevent a potential impact being effectively detected in these estimates. For example, the employment data is reported at the worksite, but specialty contractors in either of the three broad NAICS sectors we examine may work at a data center construction site for only a few weeks. In that case, employment within a treated county may not be correctly reported in the county where the work took place. That would affect Construction, and perhaps specialty workers in the Information sector, but not those in Professional and Technical Services who might be associated with data centers as permanent workers.

The most likely explanation for the issue is actual heterogeneity in data center use, and hence labor demand. The distinction between registered and large data centers in Texas occurs at a low size threshold of 100,000 square feet. But, as we mentioned above, data centers have many uses, and the demand for operating employment likely varies across types not apparent in the count data we use in these estimates.

The presence of significant heterogeneity in labor demand surrounding data centers is itself an important finding and prompts some additional analysis. The location of the most rapid expansion of data centers is Taylor County (Abilene), which now contains four total data centers, each of which arrived in 2021 or later. However, in the sectors most closely tied to the opening and operation of data centers, net employment in Taylor County has seen only small increases in jobs as a consequence of the first suite of new data centers. Figure 6 demonstrates this visually for a subset of 4-digit NAICS employment codes that are most likely to be impacted by the arrival of a new data center.

Figure 6: Taylor County (Abilene) Selected Employment During Early Data Center Opening



Summing the aggregate job growth in these four sectors before any data center entry (Q4 2020) and the latest period in our dataset (Q4 2024) implies total job increases of only approximately 320 jobs – less than 5% of the roughly 6,700 jobs Taylor County added during this period.

On aggregate, factors such as the size of the individual data center, or even clustering, seem not to play a major role in labor demand. Other factors, such as the data center end user mix, differences in equipment monitoring technology/policies, or latency (distance from data center) of users likely play a larger role in the labor demand. Absent clear data on these, much of which will likely remain invisible to the econometrician, data centers are likely to experience considerable measured heterogeneity in labor demand.

5 Summary and Discussion

The rapid expansion of digital infrastructure, particularly data centers, has significantly redirected local capital attraction efforts across the United States. As these facilities transition to central nodes of an AI-intensive economy, their influence on local labor markets has become a primary subject of policy debate. While municipal leaders frequently champion data center projects as catalysts for high-tech employment and expanded tax bases, the empirical reality of their long-term labor market returns remains ambiguous. This study addresses this gap by providing a causal analysis of the employment effects associated with data center arrivals across Texas counties from 2000 through 2024.

The empirical findings presented in this research offer a stark departure from the high-growth narratives

prevalent in Input-Output studies (see Table 1). Across nearly every econometric specification, data centers appear to require remarkably little operational labor. Although initial Two-Way Fixed Effects (TWFE) models suggested a statistically significant but economically negligible lift to the total county employment growth rate of approximately 0.08 percentage points, this result proved non-robust. Upon application of the Callaway and Sant'Anna (2021) estimator to correct for potential biases inherent in staggered treatment timing, the evidence for a positive aggregate employment growth rate effect effectively vanishes.

Sectoral decomposition into Information (NAICS 51), Professional and Technical Services (NAICS 54), and Construction (NAICS 23) reveals no statistically significant long-term job growth rate increases. Event study estimates confirm a transient (2-year) increase in construction activity during the site preparation phase. This result is unsurprising, and though our point estimates of employment level effects are lower than those reported in Table 1 (the I-O estimates),[11] we believe that is explainable for three reasons. First, as noted above, some of the Construction employment that markets allocate to data centers likely comes from other, competing construction sectors, such as residential construction. So, the gross effect provided in an Input-Output model will partially overstate net employment changes (which is what we measure). Second, the studies in Table 1 focus primarily on large data centers. These would naturally have a larger labor demand for Construction employment. Third, some of the Construction employment effect may be missed in the administrative data (Quarterly Workforce Indicators). Some construction employees may be at the data center site for periods that are too short a duration to be reported by the underlying Census Quarterly Census of Employment and Wages.

Our results hold across estimates in which we can identify the equation through the parallel trends assumption, and in both tests of employment growth and employment levels across total employment and those most logically affected by data center entrance.

Furthermore, the data suggest a failure of the clustering hypothesis. The entrance of multiple data centers did not materially differ from a single entrance. And, despite the scale of Large Data Centers exceeding 100,000 square feet, there is little evidence that these facilities serve as anchors for high-tech employment in Professional & Technical Services or Information Services.

In both the growth models and the level estimates (provided in the Appendix), the statistical certainty of total employment, Information, and Technical and Professional Services employment drops substantially after treatment. This strongly suggests heterogeneity in the employment effects of data centers, which we have not been able to capture through our size measures.

In contrast, the confidence interval for Construction employment growth shrank substantially post-treatment, suggesting a very common increase in labor demand around deployment. Importantly, this tightening of labor demand at the treatment point gives us confidence that the administrative data on data center tax applications closely matches a common point of construction for the establishments in our sample, which is important for identification.

Likewise, the smaller data centers with fewer than 100,000 square feet are unrelated to potential high-tech employment clusters. This is a finding that contrasts with the claim by McKinsey & Company (2025b) that small data centers are more likely to exhibit job creation in specialized Professional and Technical Services. The proximity-agnostic nature of modern AI training workloads appears to decouple physical infrastructure from the high-skilled human capital markets required for inference and design. Long-term duration effects indicate that as facilities age, their impact on local employment remains neutral or occasionally turns negative, suggesting a potential crowding-out effect where the extreme land and energy requirements of these facilities displace more labor-intensive industrial opportunities.

This conclusion must be viewed with some caution, given that we are still in the early days of data center construction and use. Texas is home to one of the largest data center clusters in the world, now under construction in Abilene. The early stages of that construction appear in our data and testing, but a facility of roughly 2,000,000 square feet may result in business locations in support, such as a training facility or

[11] Note that the point estimate for the construction employment (in levels) TWFE model estimate for large data centers in Appendix Table A.4 (16%) suggests an increase of 813 jobs when applied to the median number of construction jobs for data center counties in Table 2 (5,079). This is on the low end of the I-O Direct Construction Job estimates in Table 1. Our Appendix A.5 event-study point estimate is statistically insignificant but suggests roughly a 5% post-event construction employment (in levels) increase. The corresponding estimate of 253 jobs is notably lower than the Direct Construction Jobs estimate in Table 1.

centralized maintenance operations, that differ substantially from the other data centers. That has not yet materialized, but the issue will prove an important research question of the 2030's.

Still, our Callaway and Sant'Anna (2021) group effects point towards zero long-term job creation associated with the early (pre-2022) data centers in Texas. Though we have not directly assessed the subsidy and employment effects, the fiscal implications of these findings appear to be significant. With the Texas Comptroller projecting the cost of data center tax abatement programs to reach $9 billion between 2025 and 2030, the cost per job of these incentives is extraordinarily high and, in many jurisdictions, potentially infinite. These results suggest that data centers are primarily supply-side driven, with location decisions dictated by land and power availability and fiber connectivity rather than tax sensitivity. Consequently, the current approach of attracting data centers through aggressive fiscal incentives does not appear efficacious based solely upon labor market effects. While data centers expand the property tax base, they do not function as traditional industrial engines for local job creation. Future research should consider whether other economic benefits such as increased land values or property taxes occur because of these data center incentives.

References

Alvarez, F. E., Argente, D., Chow, J., and Van Patten, D. (2026). Data Centers and Local Economies in the Age of AI: A Shift-Share Approach. Working Paper 35194, National Bureau of Economic Research.

Arzumanyan, M., Calzado, E. R., Lin, N., Bahadur, V., Das, J., Ko, T. L., and Koesterke, L. (2025). Geospatial Suitability Analysis for Data Center Placement: A Case Study in Texas, USA. *Sustainable Cities and Society*, 122:106687.

Bahar, D. and Wright, G. (2026). New Evidence on Data Center Employment Effects. Technical report, Brookings Institution.

Batson, A. (2025). North America Data Center Report Year-End 2025. Technical report, Jones Lang LaSalle.

Bloom, N., Handley, K., Kurmann, A., and Luck, P. (2019). The Impact of Chinese Trade on US Employment: The Good, the Bad, and the Apocryphal. Working Paper 25447, National Bureau of Economic Research.

Boyce, M. and Whitacre, B. E. (2026). The Economics of Data Centers: Considerations for Communities. Factsheet AGEC-924, Oklahoma State University Extension.

Byrne, D., Martin, L., and Nah, C. (2022). Price Discrimination by Negotiation: A Field Experiment in Retail Electricity. *Quarterly Journal of Economics*, 137(4):2499–2537.

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, 225(2):200–230.

Chia, L., Hu, S., Wang, Q., and Mingxuan, F. (2026). The Spatial Incidence of Hyperscale Data Centers. Working Paper, National University of Singapore.

Delgado, M. S. and Florax, R. J. G. M. (2015). Difference-in-Differences Techniques for Spatial Data: Local Autocorrelation and Spatial Interaction. *Economics Letters*, 137:123–126.

Diao, C. and Parker, J. (2025). Economic Benefits of Data Centers. Technical report, Regional Economic Consulting Group.

ECONorthwest (2014). The Economic and Fiscal Impacts of Facebook's Prineville, Oregon Data Center. Technical report, ECONorthwest.

Fang, T. and Greenstein, S. (2025). Where the Cloud Rests: The Economic Geography of Data Centers. *Strategy Science*, 10(4):404–420.

Feher, A., Garcia-Appendini, E., and Mihet, R. (2025). Is AI trained on public money? Evidence from U.S. data centers. Working Paper, Swiss Finance Institute Research.

Ferreira, J., Shobe, W., Rephann, T., and Scheffel, M. (2026). Economic, Fiscal and Energy-related Impacts of Data Centers in the Great Lakes Region. Technical report, University of Virginia.

Gamache, C. K., Bailey, R., Emnett, M., Ornelas, K. K., Luong, L., Makholm, J., Patterson, D., Ralls, M. A., Rendahl, A., and Tinjum, A. (2025). Foundation of the energy law journal symposium on data center issues. *Energy Law Journal*, 46:335–364.

Gargano, A. and Giacoletti, M. (2025). Subsidizing the cloud: US state incentives to data centers. Working Paper, University of Houston.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*, 225(2):254–277.

Gramlich, J., Kennedy, B., McClain, C., and Stocking, G. (2026). How Americans view data centers' impact in key areas, from the environment to jobs. Pew Research Center.

Hicks, M. J. and Faulk, D. (2024). The Impact of a State-Level, Place-Based Economic Development Program: A Difference-in-Differences Approach. *Economic Development Quarterly*, 38(3):195–208.

Hicks, M. J., Faulk, D., Orzechowski, S., and Payton, S. (2026). Understanding the Economic Effects of Renewable Energy Restrictions: Evidence from Indiana, USA. *Energy Policy*. Forthcoming.

Joint Legislative Audit and Review Commission (2024). Data Centers in Virginia. Technical report, JLARC.

Justice, J. S. and Shepherd, T. L. (2025). Georgia Data Center Sales and Use Tax Exemption. Technical report, Carl Vinson Institute of Government, University of Georgia.

Kenney, P. (2025). Texas: The Backbone of the U.S. AI Ecosystem. Technical Report, Syntax.

Kheybari, S., Monfared, M. D., Farazmand, H., and Rezaei, J. (2020). Sustainable Location Selection of Data Centers: Developing a Multi-Criteria Set-Covering Decision-Making Methodology. *International Journal of Information Technology and Decision Making*, 19(3):741–773.

Knittel, C., Senga, J., and Wang, S. (2025). Flexible Data Centers and the Grid: Lower Costs, Higher Emissions? Working Paper No. 34065, National Bureau of Economic Research.

Kollar, A. and Grady, C. (2025). The Relationship Between Data Centers and the Climate Is a Systems Challenge: A Spatial Analysis of United States Data Centers. *Environmental Research Communications*, 7(11):111005.

Leka, E. (2022). What Should Economists Know About the Cloud? A Literature Review on Digital Economics. In *31st European Conference of the International Telecommunications Society (ITS)*, Gothenburg, Sweden.

LeRoy, G. and Tarczynska, K. (2025). Cloudy With a Loss of Spending Control: How Data Centers Are Endangering State Budgets. Technical report, Good Jobs First.

Lin, N., Arzumanyan, M., Calzado, E. R., Foreman, D., and Schuba, N. (2024). Data Center Growth in Texas: Energy, Infrastructure, and Policy Pathways. Technical report, Bureau of Economic Geology, The University of Texas at Austin.

Makholm, J. D., Olive, L. T. W., and Beiser, E. A. (2024). Hyperscale Data Centers, Regulatory Institutions, and US Economic Growth. *Climate and Energy*, 41(4):21–26.

Mangum Economics (2023). Research Note: Impact of the 2019 Illinois Data Center Incentive. Technical report, Mangum Economics.

Mangum Economics (2026). The Impact of Data Centers on Virginia's State and Local Economies. Technical report, Northern Virginia Technology Council.

McKinsey & Company (2025a). The Data Center Balance: How US States Can Navigate Opportunities and Challenges.

McKinsey & Company (2025b). The Next Big Shifts in AI Workloads and Hyperscaler Strategies.

Morrow County (2024). Economic Impact Analysis of Prospective Exascale Data Center Development in Morrow County, Oregon. Technical report, Morrow County.

National Conference of State Legislatures (2025). Policy Snapshot: Data Center Initiatives. Technical report, National Conference of State Legislatures.

Ngata, W. M., Bashir, N., Westerlaken, M., Liote, L., Chandio, Y., and Olivetti, E. (2025). The Cloud Next Door: Investigating the Environmental and Socioeconomic Strain of Datacenters on Local Communities. In *Proceedings of the ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies*, pages 769–774.

Nguyen, T. and Green, B. (2025). What Happens When Data Centers Come to Town? Technical report, Gerald R. Ford School of Public Policy, University of Michigan.

PricewaterhouseCoopers (2023a). Economic, Environmental, and Social Impacts of Data Centers in the United States. Technical report, Prepared for The Data Center Coalition.

PricewaterhouseCoopers (2023b). Economic, Environmental, and Social Impacts of Data Centers in the United States Including Statewide Impacts for Arizona, Ohio, and Virginia. Technical report, Prepared for The Data Center Coalition.

Priest, A. (2026). Not In My Back Yard! The Effects of Data Centers on Housing Prices. Working Paper, University of Rochester.

Riveros-Gavilanes, J. M. (2023). A Simple Test of Parallel Pre-trends for Differences-in-Differences. MPRA Paper No. 119367.

Roychoudhury, S. and Ross, J. M. (2026). Modeling Economic Impacts of a Hyperscale Data Center Campus. Working Paper, Capital University.

Rubinovitz, M. (2026). Here Comes the Cloud: The Effect of Data Centers on Homeowners in Northern Virginia. Working Paper, George Washington University.

Ryu, A. and Hiatt, S. R. (2025). Data Center Employment Forecast Analysis. Technical report. Hamm Institute for American Energy Oklahoma State University.

Sage Policy Group (2024). The Economic and Fiscal Impact of Three Rowan Digital Infrastructure Developments in Frederick County. Technical report, Sage Policy Group.

Siddick, M., Shehabi, A., and Marston, L. (2021). The Environmental Footprint of Data Centers in the United States. *Environmental Research Letters*, 16(6):064017.

SRC Evalmetrics (2025). Economic Impact Study of Data Centers in Ohio. Technical report, Ohio Chamber of Commerce Research Foundation.

Thompson, E. (2019). Economic and Fiscal Impacts of Sarpy County Data Centers. Technical report, Bureau of Business Research, University of Nebraska–Lincoln.

Tian, Z., Lei, Z., He, Z., Zhuang, Y., and Goetz, S. (2026). Data Centers and Local Economic Impacts: Insights from the Early Data Centers. In *Rural Economic Development Summit: Data Centers and Digital Infrastructure*, Harrisburg, Pennsylvania. Conference presentation.

U.S. Chamber of Commerce Technology Engagement Center (2017). Data Centers: Jobs and Opportunities in Communities Nationwide. Technical report, C_TEC.

Vaughan, J. (2025). U.S. Data Center Tax Incentives: A Special Report. Technical Report, Data Center Knowledge.

Waters, K. and Clower, T. (2025). Data Centers and 2023 Home Sales in Northern Virginia. Technical report, Center for Regional Analysis, George Mason University.

Yue, D. and Zeng, Y. (2026). The Local Economic Effects of Data Center Entry. Working Paper, Georgia Institute of Technology.

Appendix

Table A.1: Supplementary Summary Statistics for Total Employment in Texas Counties (2000–2024)

	Mean	Median	Maximum	Minimum	Std. Dev.
Total Employment	36,450	3,856	2,311,601	23	169,038
Construction Employment	2,605	269	184,721	0	12,092
Information Employment	953	38	85,812	0	4,975
Professional and Technical Services Employment	2,720	104	243,983	0	16,358
Total Data Centers	0.030	0	9	0	0.294

Table A.2: Stable Unit Treatment Value Assumption (SUTVA) Test

Dependent Variable: Mean Employment of Adjacent Counties ($\widehat{WY}$)	Data Centers (β)
Total Employment	-0.0001
	(0.0001)
Construction Employment	-0.005
	(0.03)
Information Services Employment	-0.013
	(0.01)
Professional & Technical Services Employment	-0.001
	(0.04)

Note: Treatment is the number of data centers in county i during quarter t. Standard errors are reported in parentheses and clustered at the county level.

Table A.3: Difference-in-Differences Estimates for Log Employment Levels (2000–2024)

	Total Employment	Construction Employment	Information Services Employment	Professional & Technical Services Employment
Data Centers (β)	0.037**	0.036	0.670	0.062*
	(0.02)	(0.03)	(0.42)	(0.03)
County FE (i)	Yes	Yes	Yes	Yes
Quarter FE (t)	Yes	Yes	Yes	Yes
N	25,387	25,007	23,195	24,606
R^2	0.98	0.94	0.92	0.95
Parallel Trend Statistic	2.99***	0.89	0.39	1.04

Note: Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023). *, **, and *** denote statistical significance at the $p < 0.10$, $p < 0.05$, and $p < 0.01$ levels, respectively.

This estimate yields results that are largely consistent with the growth rate tests reported in the main text. In the total employment estimate, we strongly reject the parallel trends assumption, so the estimate is silent regarding the potential aggregate causal employment effects of data centers. This is not apparent in the event studies, however. Neither the Construction nor Information Services coefficients rise to statistical significance. The Professional and Technical Services employment coefficient is marginally significant, but is relatively large, suggesting a data center increases employment in this sector by 6.2%.

For the Professional and Technical Services, the overall effect on employment ranges from trivial to modest in levels, as both sectors are modest in the average Texas county. In the median Texas county, the effect of a data center for Professional and Technical Services is just under 8 workers. In the mean county, the effect is roughly 184 workers.

Again, changes to employment levels are also of interest. Tables A.4 and A.5 report the Difference-in-Difference estimate of the natural log of employment in levels across these four sectors. We test the effect of

large Data Centers ($> 100,000$ square feet) and Registered ($< 100,000$ square feet) data centers.

Table A.4: Difference-in-Differences Estimates for Employment Levels from Large Data Centers in Texas ($> 100{,}000$ square feet), 2000–2024

	Total Employment	Construction Employment	Information Services Employment	Professional & Technical Services Employment
Coefficient (β)	0.120***	0.160**	0.070	0.150**
	(0.04)	(0.07)	(0.08)	(0.06)
County FE (i)	Yes	Yes	Yes	Yes
Quarter FE (t)	Yes	Yes	Yes	Yes
N	25,387	25,007	23,195	24,606
Adjusted R^2	0.99	0.94	0.93	0.95
Parallel Trend Statistic	2.81***	0.16	0.63	2.13***

Note: Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023). *, **, and *** denote statistical significance at the $p < 0.10$, $p < 0.05$, and $p < 0.01$ levels, respectively.

Table A.5: Difference-in-Differences Estimates for Employment Levels from Registered Data Centers in Texas ($< 100{,}000$ square feet), 2000–2024

	Total Employment	Construction Employment	Information Services Employment	Professional & Technical Services Employment
Coefficient (β)	0.033*	0.020	0.090	0.060
	(0.02)	(0.03)	(0.06)	(0.04)
County FE (i)	Yes	Yes	Yes	Yes
Quarter FE (t)	Yes	Yes	Yes	Yes
N	25,387	25,007	23,195	24,606
Adjusted R^2	0.99	0.94	0.94	0.95
Parallel Trend Statistic	2.69*	0.87	0.60	0.34

Note: Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023). * denotes statistical significance at the $p < 0.10$ level.

We find that employment effects appear to vary only modestly with facility scale. In Table A.4 we report a failure to identify both the total employment and Professional and Technical services employment because we reject the pre-trend assumption in both cases. In Table A.5, the registered ($< 100,000$ square feet) estimate, the assumption of parallel trends is rejected in the total employment difference in difference estimate.

Thus, across both tables, only one result can be identified, which is the construction employment effect in large data centers. In this estimate, a single new data center would raise construction employment by 16%, or by 43 workers in the median county or 416 workers in the mean Texas county. This is only modestly lower than estimates of construction jobs reported in previous studies in Table 1.

Table A.6: Disaggregated Employment Levels Difference-in-Differences Estimates (2000–2024)

	Information (NAICS 517)	Data Processing, Hosting and Related Services (NAICS 518)	Computer Systems Design (NAICS 5415)
Coefficient (β)	-0.008	0.038	-0.007
	(0.04)	(0.07)	(0.06)
County FE (i)	Yes	Yes	Yes
Quarter FE (t)	Yes	Yes	Yes
N	20,654	7,998	18,263
Adjusted R^2	0.89	0.89	0.91
Parallel Trend Statistic	1.88*	1.50	-0.54

Note: Standard errors are clustered at the county level. Parallel trends statistic derived from Riveros-Gavilanes (2023). * denotes statistical significance at the $p < 0.10$ level.

Figure A.1: Difference in Difference, Total Employment

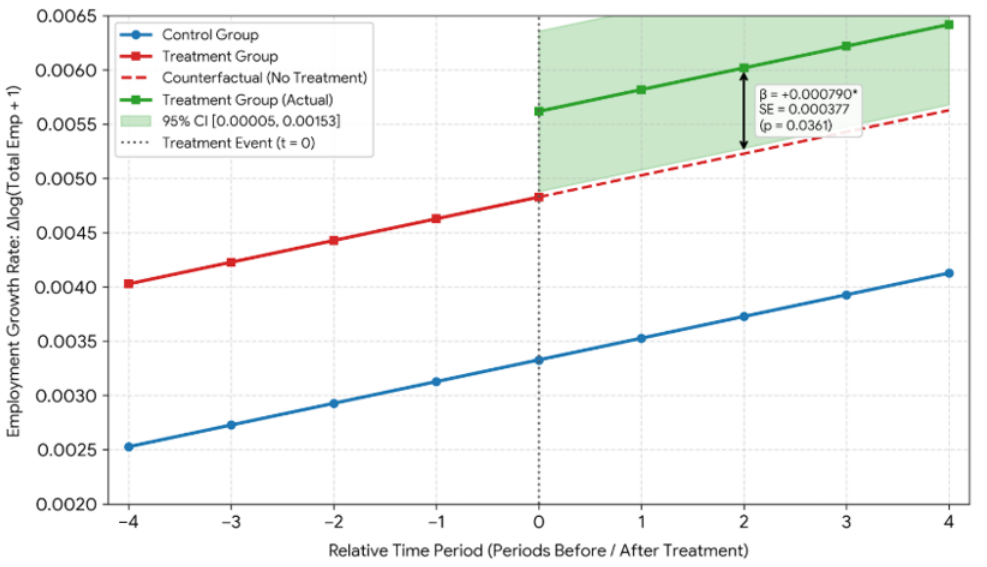


Figure A.2: Difference in Difference, Construction Employment

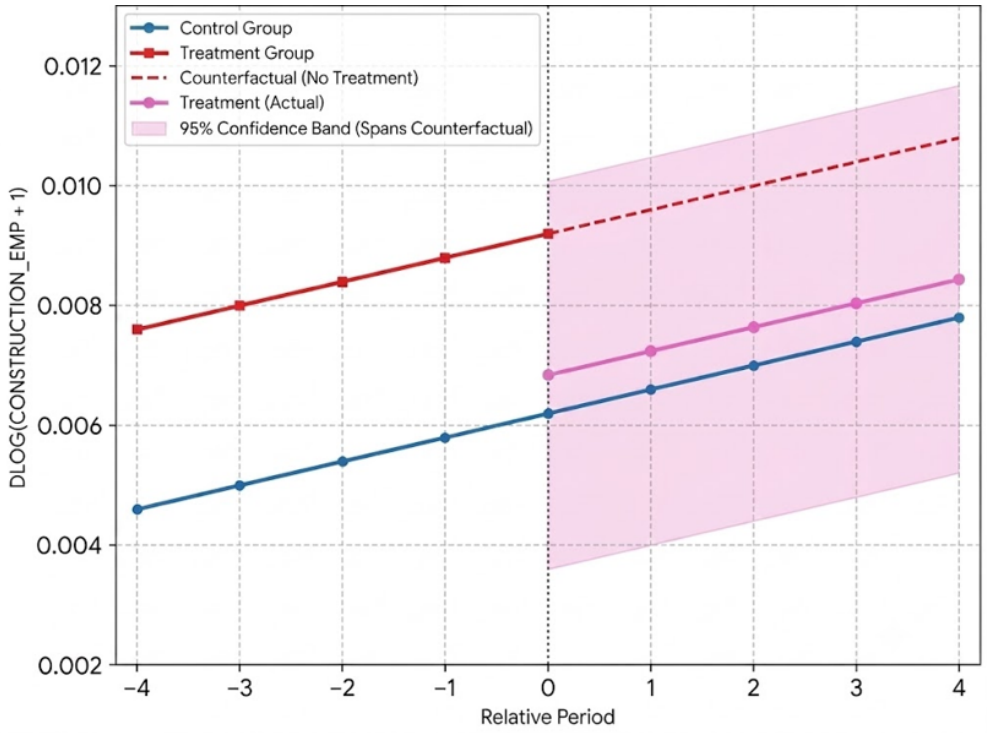


Figure A.3: Difference in Difference, Information Employment

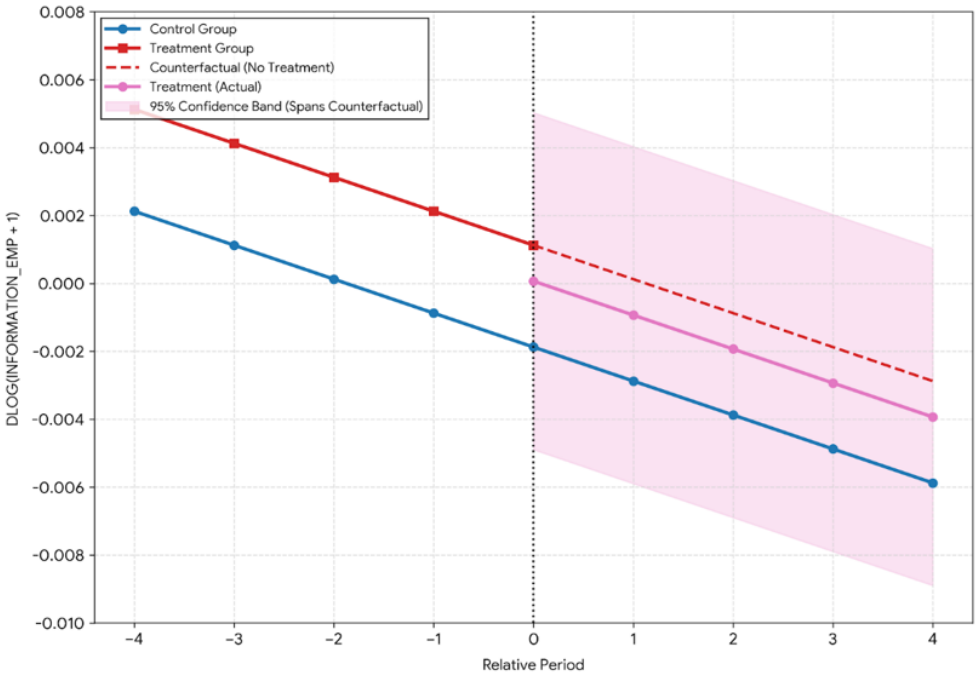


Figure A.4: Difference in Difference, Professional and Technical Services Employment

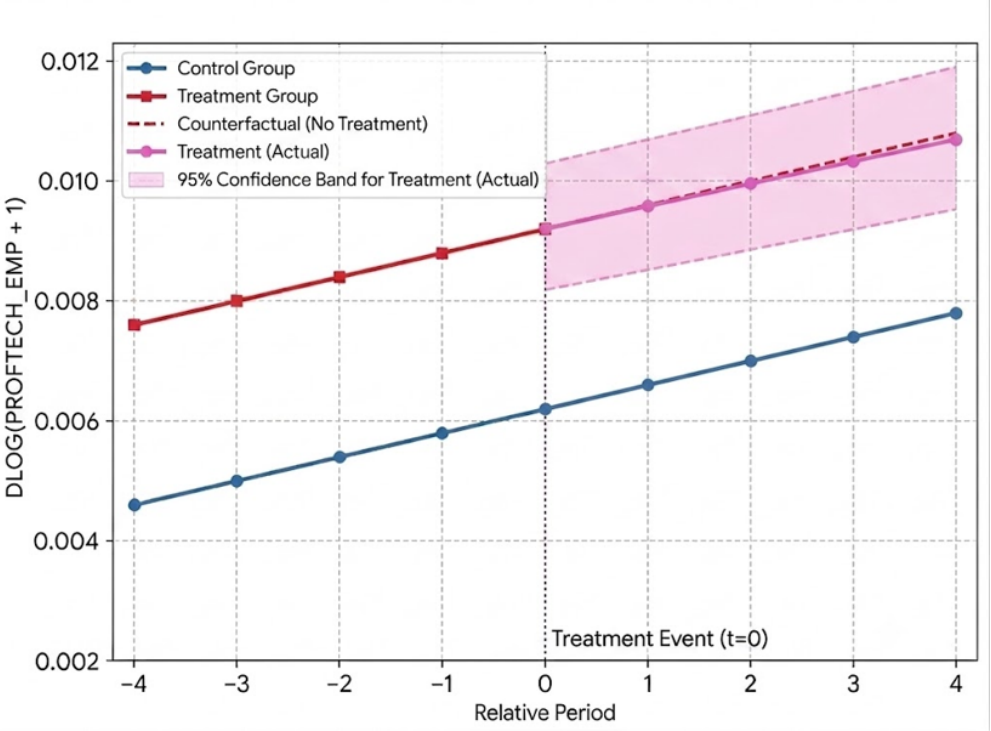


Figure A.5: Employment Effect of all Data Centers in Texas (log Levels)

