

Title

Serving the Servers: Do Data Centers Create Local Employment?

Authors

**Ernesto Aldana
Simon Büchler
Alexander Carlo
Lyndsey Rolheiser
Albert Saiz
Carles Vergara**

Date

September 2026

MIT Center for
Real Estate Research
Paper No.

26/11

Serving the Servers: Do Data Centers Create Local Employment?¹

Ernesto Aldana, Simon Büchler, Alexander Carlo, Lyndsey Rolheiser, Albert Saiz, Carles Vergara

MIT Center for Real Estate Research Paper No. 26/11

September 2026

Abstract

Data centers are among the most capital-intensive investments in the modern economy, yet they employ relatively few workers directly. Do these investments generate jobs elsewhere in the local economy? Combining proprietary data on 1,584 U.S. data centers with synthetic difference-indifferences and local projections, we find that local non-construction employment changes little for two years after a data center opens, then begins to rise. Installed capacity is associated with 7.4 additional non-construction jobs per megawatt (MW). Because each installed MW also predicts a further 1.7 MW of capacity over the same period, this corresponds to 2.74 jobs per effective MW. These gains are concentrated in two sectors: transportation/warehousing, and professional/technical services. Retail, restaurants, and other local services show little employment growth. Existing capacity mostly predicts expansion by incumbent operators, and only a small share of subsequent capacity is attributable to new entrants. Large capital investments can therefore generate local jobs even when the facilities themselves employ few workers.

Ernesto Aldana

Clemson University

Simon Büchler

Miami University

Alexander Carlo

York University

Lyndsey Rolheiser

York University

Albert Saiz

Massachusetts Institute of Technology

Carles Vergara

IESE Business School

The views expressed herein are those of the authors and do not necessarily reflect the views of the MIT Center for Real Estate.

MIT Center for Real Estate working papers are circulated for discussion and comment purposes and have not been peer-reviewed.

¹We thank participants at the 2026 AREUEA International Conference for helpful comments and suggestions. We also thank Román Modrego and Alp Oenguec for excellent research assistance..

1 Introduction

Data centers have become one of the largest categories of infrastructure investment. The rapid expansion of cloud computing services and artificial intelligence (AI) has sharply increased demand for computing capacity. Recent estimates suggest that U.S. data center investment will reach \$370 billion in 2026, while nearly \$7 trillion could be invested worldwide by 2030.¹ Although the computing services supported by data centers can be accessed from virtually anywhere, the facilities themselves are large, place-based capital investments whose effects on surrounding economies remain poorly understood.

The local stakes are substantial. Federal initiatives seek to accelerate the development of AI infrastructure (White House, 2025a,b). At the same time, state and local governments compete to attract data centers with tax incentives and subsidies (National Conference of State Legislatures, 2026; Gargano and Giacoletti, 2025).² But the local economic case for these incentives is not obvious. Data centers require large amounts of electricity and water (Mytton, 2021), can raise electricity prices (Kay et al., 2026), and generate other local environmental costs (Gour et al., 2026). Once construction ends, they directly employ relatively few workers (Ryu and Hiatt, 2025). These tradeoffs have fueled local opposition to new projects (Faverio and Kikuchi, 2026; Mathur and Sharma, 2026).³ The economic case for attracting data centers therefore depends in part on whether these investments generate broader local economic activity.

Data center investment can affect local economies through several channels.⁴ Data centers rely on local suppliers for maintenance, security, real estate, and other services (Sarte et al., 2024). The resulting employment and income can also increase demand for restaurants, retail, and other local services (Carlino and Saiz, 2019; Almagro and Domínguez-Iino, 2025). Data centers may also generate within-industry agglomeration economies. A new facility can benefit from the fiber networks, power infrastructure, and connectivity already built around existing facilities. Whether

¹Brandsaas et al. (2025) forecast annualized investment of approximately \$370 billion by the second quarter of 2026. McKinsey estimates that global capital expenditures on data center infrastructure could reach nearly \$7 trillion by 2030, with more than 40 percent occurring in the United States.

²For example, Texas provides sales tax exemptions on data center equipment and electricity conditional on investment, employment, and facility-size requirements.

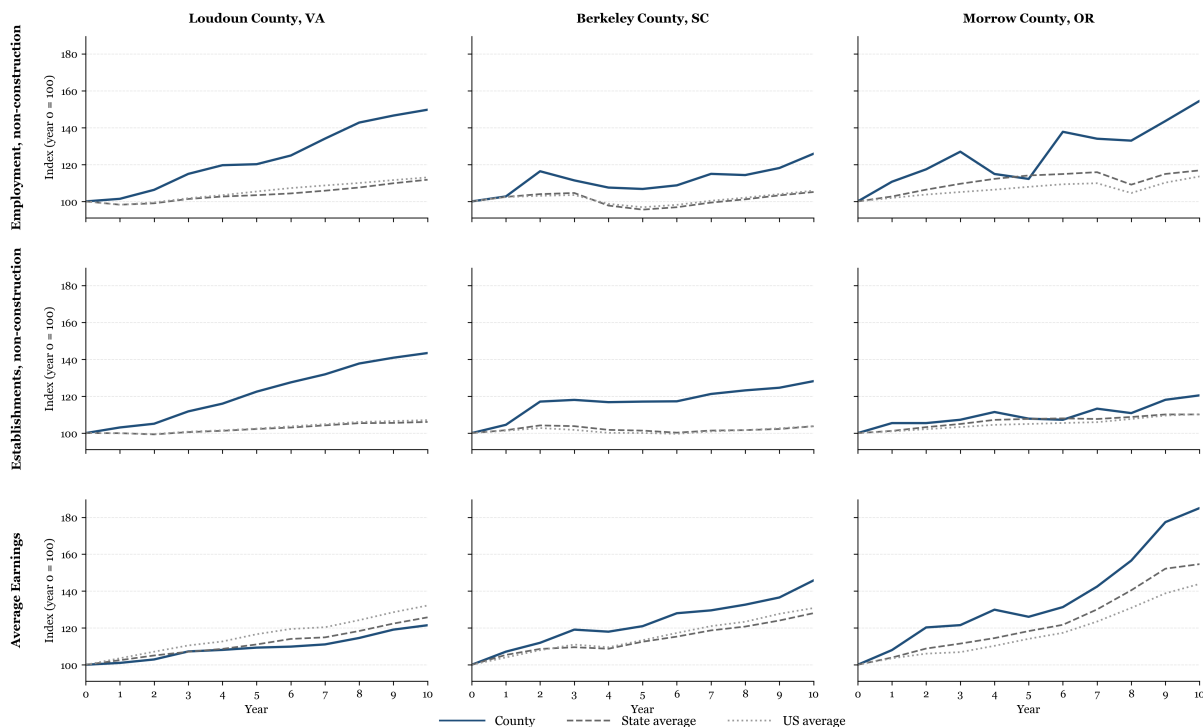
³According to Data Center Watch, local opposition contributed to the delay or cancellation of data center projects worth approximately \$156 billion in 2025; see *The New York Times*, March 26, 2026. See also Mathilda Scott, “Updates on the Cloud: More Moratoriums on Data Centers,” Rockefeller Institute of Government, June 10, 2026, which documents recent state and local efforts to restrict or temporarily pause data center development.

⁴The computing services supported by data centers, including AI, may have large national-level effects on labor and capital productivity across sectors (Acemoglu and Restrepo, 2018). We focus instead on the localized economic effects of the facilities themselves.

existing data center activity attracts new operators is ultimately an empirical question.

Figure 1 provides a first look at local economic activity around major data center development. We focus on three different markets: Loudoun County, Virginia; Berkeley County, South Carolina; and Morrow County, Oregon. Loudoun is the country’s largest data center market, while Berkeley and Morrow illustrate the industry’s expansion into smaller local economies. The figure traces non-construction employment,⁵ establishments, and average earnings around data center development. All three counties experienced substantial changes in local economic activity over this period. But these patterns are descriptive. They do not tell us how these counties would have evolved without data center investment.

Figure 1: **Local Economic Activity Around Data Center Development: Three Illustrative Counties**



Note: The figure shows non-construction employment, establishments, and average earnings around data center development in Loudoun County, Virginia; Berkeley County, South Carolina; and Morrow County, Oregon. For Berkeley and Morrow Counties, year 0 marks the first observed data center activation, in 2008 and 2015, respectively. Because Loudoun County already hosted a data center at the beginning of the sample, year 0 instead marks the first observed activation above 50 MW, which occurred in 2000 and added 120 MW of capacity. Outcomes are indexed to 100 in year 0. Solid lines represent county outcomes, long-dashed lines the corresponding state averages, and dotted lines the U.S. average. State and national averages are shown for comparison only.

⁵We exclude construction employment because it largely reflects temporary labor demand associated with building the facility.

Data center location is endogenous.⁶ Electricity costs, fiber connectivity, available land, tax incentives, and other local conditions shape where facilities are built (Du et al., 2026; Carlo and Rolheiser, 2026). These conditions can also shape local employment growth. A county that never hosts a data center may therefore be a poor comparison for one that does. Scale creates a different problem. Most U.S. counties have no installed data center capacity.⁷ Among counties with data centers, installed capacity varies widely.⁸ County labor markets also differ greatly in size. The same percentage change can therefore represent very different changes in capacity or employment, making conventional logarithmic specifications difficult to interpret (Chen and Roth, 2024).

Investment continues after initial entry, and existing facilities frequently expand. An event study around first activation captures these later investments as part of the average post-entry response across heterogeneous facilities. In addition, it cannot tell us how the response varies with installed capacity. The opposite problem arises when each facility’s entry or expansion is treated as a separate event because it ignores the capacity already operating.

Our data let us observe when data centers become operational, where they are located, and their installed computing capacity, including subsequent expansions. The proprietary Aterio database covers U.S. data centers from 2000 through 2023.⁹ Our analysis follows 1,584 active facilities with known activation dates and links them to county-level employment, establishments, earnings, and other local outcomes. We use the timing of the first data center activation to estimate the average effect of the initial entry across counties (the extensive margin of entry). Synthetic difference-in-differences compares treated counties with never-treated counties that had similar employment paths before activation. We then use installed capacity to study how local outcomes vary with the scale of data center investment within counties (an intensive margin). Local projections relate cumulative changes in these outcomes to installed MW across counties and over time. Because we estimate the specifications in levels, the employment coefficients are measured in jobs per MW.

The results indicate a gradual increase in local employment following data center entry. Non-construction employment increases slowly in the first two years after activation. It then rises by roughly 1,000 to 1,300 jobs at horizons of three through five years, although the estimates are

⁶Appendix A.1 discusses in detail the empirical challenges that arise when estimating the local effects of data center investment and how they shape our empirical approach.

⁷Of the 3,121 counties included in the geographic sample, 2,860 (91.6%) had no active data center by 2023.

⁸Data center capacity is conventionally measured in megawatts (MW), capturing the maximum instantaneous power available to operate servers, cooling systems, and supporting infrastructure. Modern data center racks vary substantially in density and computing power, making MW a more informative measure of capacity than square footage (Shehabi et al., 2024).

⁹Our sample therefore predates much of the recent AI-driven expansion in hyperscale data center investment. Our estimates should not be mechanically extrapolated to the newest generation of AI-focused facilities.

imprecise. The intensive-margin results show a similar buildup over time. One additional MW of installed capacity is associated with 7.4 additional non-construction jobs after five years and predicts 1.7 MW of subsequent capacity. Accounting for this additional capacity, the employment response corresponds to 2.74 jobs per effective MW.

The employment gains are sectorally concentrated. Transportation and warehousing and professional and technical services show the largest employment responses. House prices increase in some specifications. Retail, accommodation and food services, information, and administrative services show little evidence of employment growth. Our capacity results point to persistence around established data center locations. Most follow-on investment reflects continued investment by operators already present in the county. However, incumbent investment also predicts subsequent capacity from new operators, which accounts for about 30 percent of the associated capacity growth. In contrast, new-operator entry does not predict successive entry by additional operators. Establishment estimates are positive but imprecise, while earnings increase modestly, particularly outside Loudoun County. The main employment patterns are robust to alternative estimators, employment data sources, and excluding the largest data center hubs.

The composition of employment growth suggests that production linkages play an important role: employment rises most in sectors that support data center operation, while local service sectors respond much less. Combined with the persistence of investment around established locations, these patterns are consistent with local supplier demand and established data center clusters mattering more than broad local expenditure effects.

Our findings contribute to two strands of research. First, we add to the emerging literature on the local economic effects of data centers. Existing studies find limited employment effects at the state level (Hicks, 2025; O’Leary, 2025) and mixed effects at finer geographic scales (Gargano and Giacoletti, 2025; Alvarez et al., 2026; Bahar and Wright, 2026). We study both the arrival and scale of data center investment. This distinction lets us trace employment after first activation and quantify how local employment varies with installed MW. Our empirical approach builds on the urban economics literature, which shows that local growth is partly predictable from initial conditions. We also address several outstanding empirical issues in innovative ways (see Appendix A.1).

Second, our findings speak to the broader literature on agglomeration and place-based investment. Local suppliers, shared infrastructure, thick labor markets, and local demand for non-tradable goods can shape how large investments affect nearby economic activity (Marshall, 1920; Rosenthal and Strange, 2004; Duranton and Puga, 2004; Greenstone et al., 2010; Kline and Moretti, 2014;

Baum-Snow et al., 2024). These effects can be empirically substantial (Greenstone et al., 2010; Kline and Moretti, 2014). Data centers provide an unusual setting. They require enormous fixed investments but relatively little direct labor. They can deliver output remotely, so customers need not locate near the facilities. Our results show that even this type of capital-intensive investment can be associated with substantial employment growth where the capital is installed.

The rest of the paper is organized as follows. Section 2 develops the conceptual framework, Section 3 describes the data, Section 4 presents the empirical analysis, and Section 5 concludes.

2 Conceptual Framework

Data center investment can affect local economies through three channels: direct employment, demand for local suppliers and services, and subsequent data center investment. Let K denote data center computing capacity, measured in megawatts (MW). Direct data center employment can be written as

$$L_D = \lambda(K)K, \quad (1)$$

where $\lambda(K)$ is direct employment per unit of installed capacity. Data centers are labor-light when $\lambda(K)$ is small. Industry calculations and our own evidence suggest that employment intensity declines with facility scale, so that $\lambda'(K) < 0$.¹⁰ Larger facilities can therefore add substantial computing capacity without a proportional increase in direct employment.

Data centers also rely on local suppliers for inputs and services. We denote these supplier sectors by S . We denote retail, restaurants, and other local service sectors by N . Letting a hat denote a proportional change, employment in either sector can respond to changes in data center capacity according to

$$\widehat{L}_j = \theta_j \widehat{K}, \quad j \in \{S, N\}. \quad (2)$$

A positive θ_S is consistent with production linkages through demand for complementary local inputs and services. A positive θ_N is consistent with additional employment and income increasing demand for local goods and services. The sectoral composition of employment growth therefore helps distinguish these two channels.

Existing data center capacity may also predict subsequent data center investment. We separate

¹⁰Ryu and Hiatt (2025) document that data center operational employment intensity per MW declines with facility capacity. Highly automated hyperscale facilities require fewer full-time operations staff per MW than smaller enterprise facilities. These smaller facilities rely on more manual, less automated operations with greater on-site staffing. This motivates modeling λ as decreasing in K .

subsequent investment into expansion by incumbent operators and entry by new operators:

$$K_{t+1} = K_t + \Delta K_{t+1}^I + \Delta K_{t+1}^E, \quad (3)$$

where ΔK_{t+1}^I denotes additional capacity from incumbent operators and ΔK_{t+1}^E denotes capacity from new operators. Shared infrastructure, network connectivity, and other local advantages may make established markets attractive for further investment. Incumbent expansion, however, may also reflect phased development of an existing campus or persistent site advantages. New-operator entry provides stronger evidence that established data center locations offer advantages that extend beyond the incumbent operator. We therefore examine the two margins separately.

These channels yield different empirical patterns. Production linkages imply employment growth in sectors that supply data centers. A local expenditure channel implies growth in retail, restaurants, and other local services. Colocation implies that existing capacity predicts subsequent investment, particularly by new operators. Appendix A.2 develops these relationships in a local equilibrium model.

3 Data and Descriptive Evidence

3.1 Data Center Location, Size, and Timing

We use proprietary asset-level data from Aterio covering U.S. data centers from 2000 to 2023. For each facility, we observe its geographic location, owner and operator, development stage, activation date, site and building size, and installed power capacity. Aterio tracks projects from announcement through construction and activation, as well as projects that are delayed or withdrawn. A project enters the database at announcement only after reaching a site-specific milestone, such as an agreement with a state or local government or the start of a formal permitting process. We do not include general expansion plans without a defined site. We define activation as the point at which a facility becomes operational. This distinction matters for timing. Construction and other investment-related activity can begin before activation, while ongoing operating employment begins once the facility becomes operational.

Table 1 provides summary statistics. Panel A reports the full database as of May 21, 2026, which includes 6,858 U.S. projects across all development stages. Of these, 1,975 data centers are operational, with a combined capacity of 45.9 GW. The development pipeline is considerably larger.

Another 3,433 projects are classified as announced, representing 301.3 GW of planned capacity. Announced facilities average 97.6 MW, compared with 24.6 MW among active facilities, showing that projects in the current pipeline are substantially larger than existing facilities. The increasing scale of new projects also motivates our focus on installed capacity rather than facility counts.

Activation dates are available for 1,967 of the 1,975 active facilities (99.6%).¹¹ Panel B reports the sample used in our analysis, which is restricted to active facilities activated between 2000 and 2023. This leaves 1,584 data centers with a combined capacity of 28.9 GW.

Table 1: Data Center Descriptive Statistics

Stage	# DCs	Total Capacity (GW)	Capacity per Asset (MW)		
			Min	Avg	Max
Panel A: Full inventory as of May 21, 2026					
All stages	6,858	445.6	0.015	71.4	900.0
Active	1,975	45.9	0.015	24.6	350.0
Announcement	3,433	301.3	0.5	97.6	700.0
Cancelled	19	1.0	20.0	56.9	200.0
Construction	793	61.9	5.0	80.8	600.0
Delayed	79	6.5	9.0	82.8	167.0
Land Bank	52	2.3	4.0	63.1	519.0
Not Approved/Withdrawn	507	26.6	25.0	69.1	900.0
Panel B: Analysis sample, activation dates through 2023					
Active	1,584	28.9	0.015	19.4	280.0

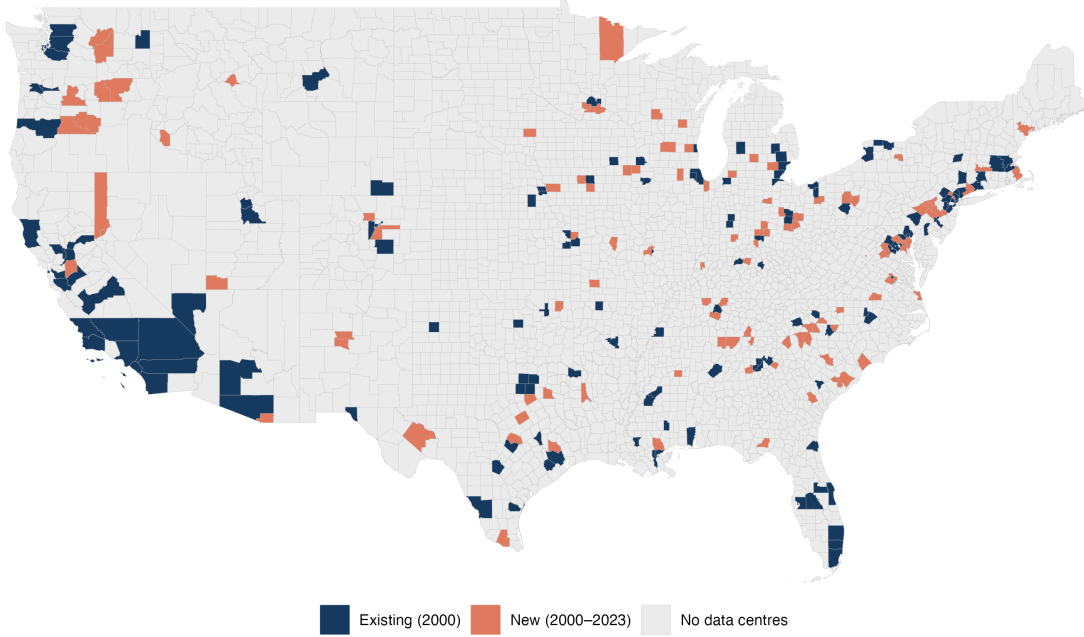
Note: The table reports the number and capacity of U.S. data center projects by development stage. Total capacity is reported in gigawatts (GW), and capacity per facility is reported in megawatts (MW). Panel A reports the full inventory observed when we accessed the data on May 21, 2026. Panel B reports the sample used in the empirical analysis, restricting projects to those activated through 2023.

Figure 2 shows how the geography of data centers changed between 2000 and 2023. Early data centers were concentrated in California and the Northeast. Since 2000, data center development has spread across much of the country, including the South, Midwest, and Mountain West. The industry nevertheless remains concentrated in a relatively small number of counties. Most U.S. counties host no data centers, while installed capacity remains concentrated in established hubs. This combination of geographic expansion and concentration generates large differences across counties in the timing and scale of data center investment.¹²

¹¹For 153 facilities, activation years were identified through manual research conducted by student research assistants and added to the Aterio data.

¹²Appendix Figures A1 and A2 provide a closer view of this geographic expansion. They map county-level data center activity in the states with the largest increases in data center counts and installed capacity.

Figure 2: **Geographic Distribution of U.S. Data Centers, 2000–2023**



Note: The figure classifies counties by the presence of active data centers. “Existing (2000)” denotes counties with at least one active data center in 2000 that remains active in 2023. “New (2000–2023)” denotes counties with no active data center in 2000 and at least one active data center by 2023. “No data centers” denotes counties without an active data center in either year.

We merge the Aterio data with county-level economic and demographic data to construct an annual county panel. The underlying data cover 2000 through 2023, while the estimating sample begins in 2001 because the specifications use outcomes or treatment changes relative to the preceding year. We consolidate Virginia independent cities into their surrounding county equivalents. Our baseline sample excludes the top 5% of counties by employment in 2000. In these large labor markets, the employment associated with a data center is too small relative to normal year-to-year fluctuations in employment, making employment-level effects difficult to detect.

3.2 Local Economic Outcomes and County Characteristics

3.2.1 Local Economic Indicators

County Business Patterns (CBP). Our primary source for local employment, wages, and business activity (establishments) is the Census Bureau’s County Business Patterns (CBP). CBP reports annual employment and establishment counts by county and industry from 2000 to 2023. We separate construction employment (NAICS 23) from non-construction employment. We also separately

examine employment in utilities (NAICS 22), retail trade (NAICS 44–45), transportation and warehousing (NAICS 48–49), information (NAICS 51), real estate (NAICS 53), professional, scientific, and technical services (NAICS 54), administrative services (NAICS 56), and accommodation and food services (NAICS 72). These industries include sectors linked to data center investment and operations, as well as nontradable services that may respond to increased local income and demand.

CBP suppresses employment in some county-industry cells to protect confidentiality. We recover these observations using the imputation procedure developed by [Eckert et al. \(2020\)](#). The procedure combines the employment ranges reported for suppressed cells with adding-up constraints across industries and geographies. The resulting panel provides our main employment measures, which we also compare with Quarterly Workforce Indicator-based measures. Historical CBP data for 1980 and 1990 allow us to measure county employment growth from 1980 to 2000.

Quarterly Workforce Indicators (QWI). We use the Census Bureau’s Quarterly Workforce Indicators (QWI) as an alternative source of employment and earnings ([U.S. Census Bureau, 2026](#)). QWI draws primarily on state unemployment insurance wage records linked to worker and firm characteristics and reports employment by county and industry. We aggregate the quarterly data to annual frequency and construct the same non-construction employment measure as in CBP. The QWI panel is shorter than CBP and contains more suppressed observations in thin labor markets.

House Price Index. We measure local house prices using the Federal Housing Finance Agency’s county-level House Price Index (HPI) ([Federal Housing Finance Agency, 2026](#)).¹³

3.2.2 Electricity and Internet Infrastructure

Electricity prices. We measure county-level industrial electricity prices using utility-level data from the U.S. Energy Information Administration (EIA) ([U.S. Energy Information Administration, 2026](#)). Electricity costs are an important determinant of data center location because these facilities consume large amounts of power ([International Energy Agency, 2025](#)). We aggregate utility prices to the county level and use state-level electricity price growth from the State Energy Data System (SEDS) to extend the series to years without county-level coverage.¹⁴ In the empirical specifications,

¹³For consolidated Virginia counties with multiple HPI series, we retain the series with the longest coverage and use 2000 population as a tiebreaker. The house price index is not available for all counties in all years, as some counties may not meet the minimum number of transactions required to produce the index.

¹⁴County-level EIA prices are available for 2011 and 2013–2023. For 2000–2010, we backcast 2011 county prices using the corresponding state’s annual electricity-price growth. We use the same procedure to impute 2012 and other missing county-year observations.

we use the county’s 2000 industrial electricity price as a baseline characteristic and interact it with a linear year trend. We do not control for contemporaneous electricity prices because of potential endogeneity.

Internet infrastructure. We measure pre-existing internet infrastructure using retrospective subnet data from CAIDA ([CAIDA, 2026](#)). The data record the number of active subnets announced at each location. We use this measure to capture baseline differences in network connectivity across counties. Established network infrastructure can make a location more attractive for data center investment.

3.2.3 County Characteristics

Baseline demographics. We measure baseline county demographics using the 2000 Decennial Census. Following the urban growth literature, the variables include the share of residents aged 25 and older with a bachelor’s degree or higher, the share of foreign-born residents, and the share of residents under age 25. When the underlying data are available only at the census-tract level, we aggregate them to the county level using population weights.

Population. We obtain annual county population from the Bureau of Economic Analysis Regional Economic Accounts ([U.S. Bureau of Economic Analysis, 2026](#)). We use population to construct the per-capita scenic amenity measure described below.

Natural amenities. We measure natural amenities using the USDA Economic Research Service Natural Amenities Scale ([U.S. Department of Agriculture, Economic Research Service, 2025](#)). The measures include standardized January temperature and the share of county land area covered by water.

Scenic amenity. We measure scenic and natural attractiveness using the number of geotagged photographs uploaded to Panoramio as of 2010, following [Saiz et al. \(2018\)](#) and [Carlino and Saiz \(2019\)](#). We normalize the measure by county population.

Presidential election returns. We measure local political preferences using county-level presidential election returns from the MIT Election Data and Science Lab ([MIT Election Data and](#)

Science Lab, 2026). Republican vote share provides a time-varying measure of county political preferences.¹⁵

4 Empirical Specifications and Results

Data centers do not locate randomly. Operators choose locations based on power infrastructure, land availability, network connectivity, and other local conditions that may also shape economic growth (Du et al., 2026; Carlo and Rolheiser, 2026). We use two complementary empirical designs to study these investments.

Section 4.1 studies the extensive margin. We use synthetic difference-in-differences (SDiD) to estimate how local outcomes change following a county’s first data center activation. The method constructs counterfactuals from never-treated counties with similar pre-treatment employment trajectories. We additionally consider a robustness exercise using local projections differences-in-differences (LP-DiD) in Appendix Section A.4.3.

Section 4.2 studies the intensive margin. We use local projections to relate changes in local outcomes to installed data center capacity, measured in megawatts (MW), across counties and over time. This design captures how local outcomes vary with the scale of data center investment and allows us to express the employment response per MW.

4.1 Extensive Margin: First Data Center Activation

4.1.1 Specification

We estimate the effect of a county’s first data center activation using synthetic difference-in-differences (SDiD) (Arkhangelsky et al., 2021) and its event-study extension (Ciccia, 2024). Data centers locate selectively, so never-treated counties may not provide appropriate counterfactuals for treated counties. SDiD addresses this concern by placing greater weight on control counties with similar pre-treatment employment trajectories. It also weights pre-treatment periods according to their relevance for constructing the counterfactual. The method extends synthetic control approaches (Abadie, 2021) to settings with multiple treated units and periods.

Treatment begins in the year a county’s first data center becomes operational and remains active thereafter. The donor pool consists of counties that do not receive a data center during the sample

¹⁵Presidential elections occur every four years. For non-election years, we assign the Republican vote share from the most recent presidential election. We exclude Alaska and Hawaii from the analysis.

period. As throughout our baseline analysis, we exclude the top 5% of counties by employment in 2000. We estimate dynamic treatment effects around the year of first activation.

Before constructing the synthetic comparison group, we residualize each outcome with respect to observed county characteristics and fixed effects:

$$\tilde{Y}_{c,t} = Y_{c,t} - \hat{\delta}_{st} - \hat{\omega}_c - \mathbf{x}'_{c,t}\hat{\gamma}, \quad (4)$$

where δ_{st} denotes state-by-year fixed effects and ω_c county fixed effects. The vector $\mathbf{x}_{c,t}$ includes Republican vote share and baseline county characteristics interacted with linear year trends. The baseline characteristics are measured in or prior to 2000 and include industrial electricity prices, internet infrastructure, employment growth from 1980 to 2000, educational attainment, the shares of foreign-born and young residents, water coverage, January temperature, and scenic amenity. These controls allow counties with different initial characteristics to follow different trends, while avoiding endogenous controls.

We then apply SDiD to the residualized outcomes. For non-construction employment, our primary outcome, we estimate dynamic treatment effects following [Ciccia \(2024\)](#). We calculate bootstrap standard errors, clustered at the county level, using Algorithm 2 of [Arkhangelsky et al. \(2021\)](#).

For all secondary outcomes, we hold the SDiD unit and time weights fixed at those estimated for non-construction employment.¹⁶ The same synthetic comparison group therefore underlies the sectoral employment, establishment, earnings, electricity price, and house price estimates. We report jackknife confidence intervals for these outcomes.¹⁷

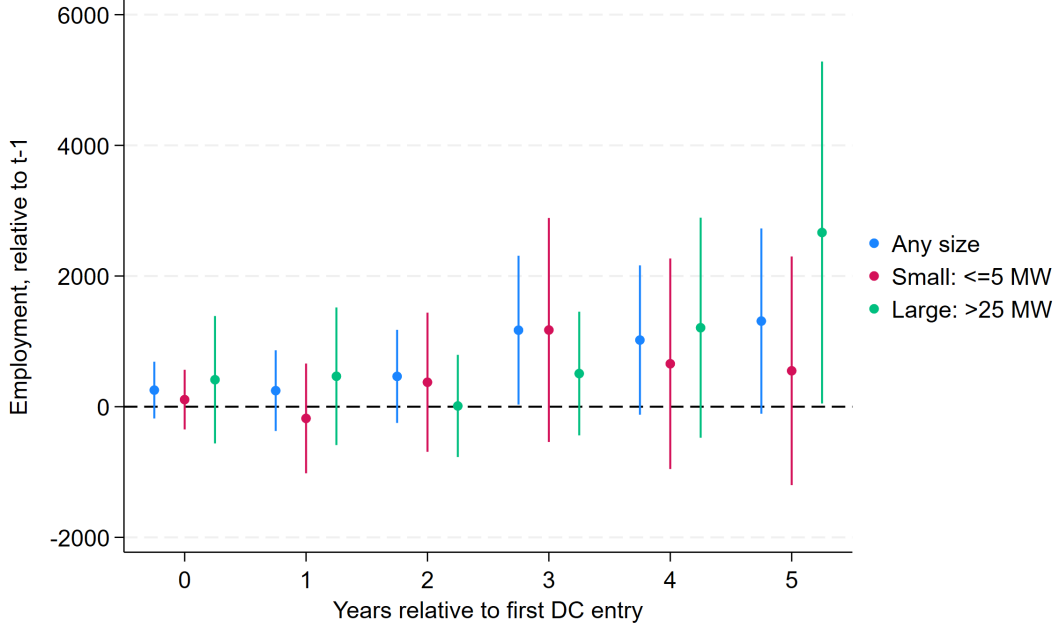
4.1.2 Employment Effects of Data Center Entry

Figure 3 shows the employment response following first data center activation. Non-construction employment changes little during the first two years. The response becomes larger after three years. At horizons three through five, treated counties have, on average, about 1,000–1,300 more non-construction jobs than their synthetic controls. The estimates are imprecise, but the pattern points to a gradual employment response.

¹⁶More precisely, we use the unit and time weights obtained from steps 1–3 of Algorithm 1 in [Arkhangelsky et al. \(2021\)](#).

¹⁷Jackknife variance estimation requires at least two estimates of sectoral employment, establishment, earnings, electricity price, and house price per county per cohort. We therefore exclude two counties treated in 2002 and 2007, respectively, when reporting these estimates. We opt for jackknife inference for this set of outcomes because, unlike with bootstrap, the SDiD weights remain constant across iterations.

Figure 3: **Employment Effects of First Data Center Activation**



Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates for non-construction employment following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The blue series includes first activations of any size. The red series restricts treatment to first activations with capacity of 5 MW or less, and the green series to first activations with capacity greater than 25 MW. The donor pool consists of counties that never receive a data center during the sample period. SDiD unit and time weights are estimated separately for each treatment definition. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Vertical bars report 95% bootstrap confidence intervals.

The delayed response may reflect several forms of local adjustment. Firms may take time to respond to new demand, workers may relocate gradually, and initial investment may be followed by additional data center capacity. Section 4.2 examines the last channel directly.

Figure 3 also separates counties by the capacity of their first facility. The red series restricts treatment to facilities of 5 MW or less, while the green series focuses on facilities larger than 25 MW. The small-facility estimates remain statistically insignificant, with point estimates of around 600 jobs at longer horizons. Large facilities show positive point estimates at most horizons, with a larger and statistically significant response after five years. Confidence intervals are wide for both groups.

These splits are substantially less precise than the pooled estimates, which include 102 counties, compared with only 55 counties in the small-facility group and 30 in the large-facility group; the remaining 17 counties, whose first facility falls between 5 and 25 MW, are not shown separately.

We therefore use first activation, regardless of size, for the remaining extensive-margin analysis.

4.1.3 Sectoral Employment Effects

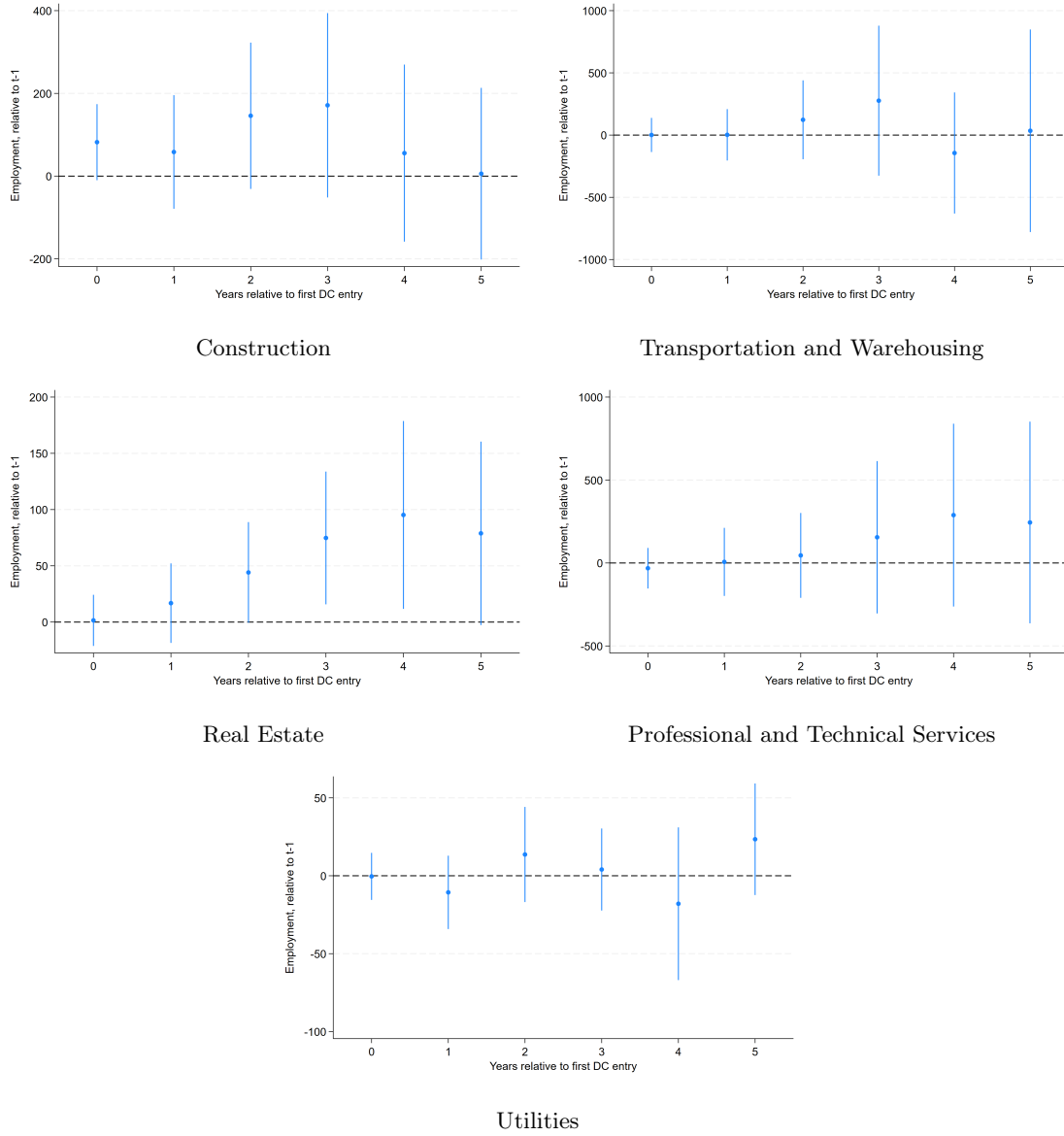
The aggregate employment response masks substantial differences across sectors. We therefore estimate effects separately by industry. To keep the counterfactual comparable across outcomes, we use the SDiD unit and time weights estimated for non-construction employment in Figure 3.

We first examine sectors that may provide inputs or services to data centers. Figure 4 reports results for construction, transportation and warehousing, real estate, professional and technical services, and utilities.

Construction follows a different timing pattern because facilities are built before they become operational. We therefore shift construction employment back three years relative to first data center activation. The estimates show construction employment rising before activation and then dissipating once the facility becomes operational. The estimates remain imprecise throughout.

Real estate shows the clearest post-activation response. Employment increases gradually and reaches approximately 100 additional jobs four years after activation. Professional and technical services also show positive point estimates at longer horizons, although these estimates remain imprecise. We find little evidence of employment growth in transportation and warehousing or utilities.

Figure 4: **Employment Effects of First Data Center Activation by Selected Sector**



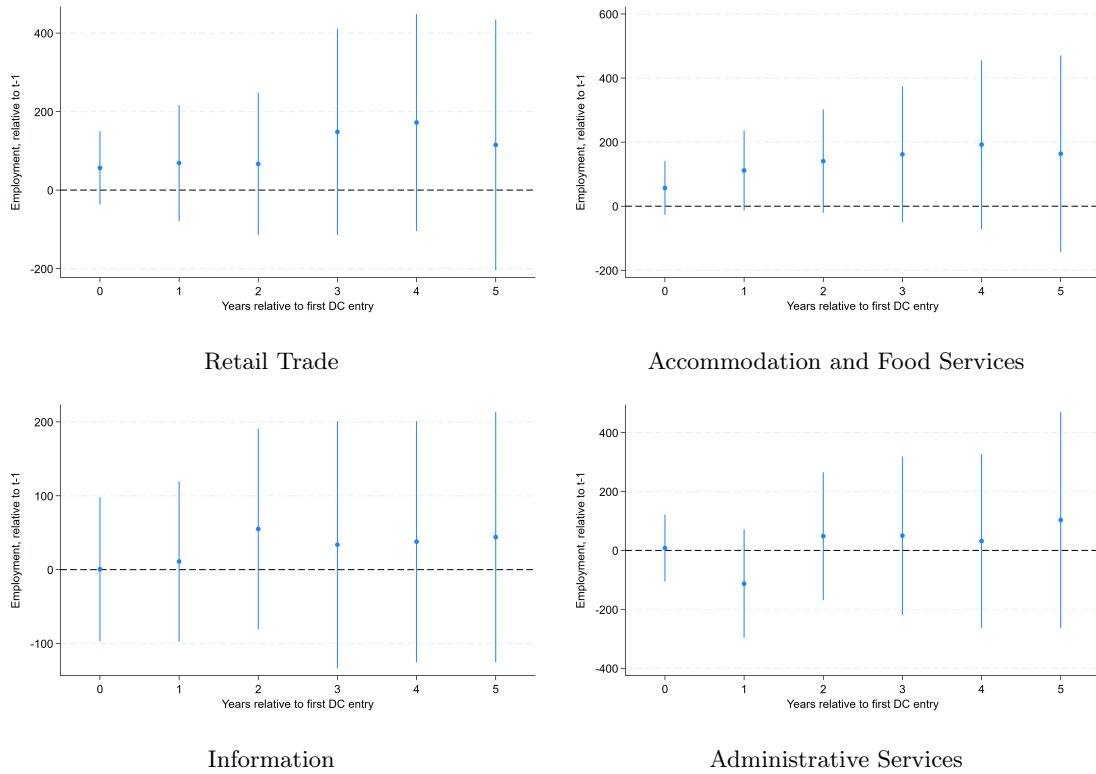
Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates for sector employment following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. For construction, employment is shifted back by 3 years relative to the first data center activation to capture activity associated with building the facility before it becomes operational. For all other sectors, event time is defined relative to the year of activation. The donor pool consists of counties that never receive a data center during the sample period. The same SDiD unit and time weights estimated for non-construction employment are used for all sector outcomes. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Vertical bars report 95% jackknife confidence intervals.

Figure 5 examines retail trade, accommodation and food services, information, and administrative services. Retail trade, information, and administrative services show no clear employment

response. Accommodation and food services show positive point estimates at longer horizons, reaching roughly 200 additional jobs four years after activation. The confidence intervals remain wide. We therefore find little precise evidence of a broad employment response in local consumer-service sectors.¹⁸

Overall, the extensive-margin estimates show employment gains in a limited set of industries. Real estate provides the clearest evidence, while professional and technical services also show positive but imprecise estimates. We find less evidence of a broad response across local consumer-service sectors.

Figure 5: **Employment Effects of First Data Center Activation in Other Local Sectors**



Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates for sector employment following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The donor pool consists of counties that never receive a data center during the sample period. The same SDiD unit and time weights estimated for non-construction employment are used for all sector outcomes. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Vertical bars report 95% jackknife confidence intervals.

¹⁸Appendix Figure A3 reports additional sectoral results. Telecommunications, computing infrastructure, and data processing show positive employment responses at longer horizons, while estimates for the remaining sectors are generally imprecise.

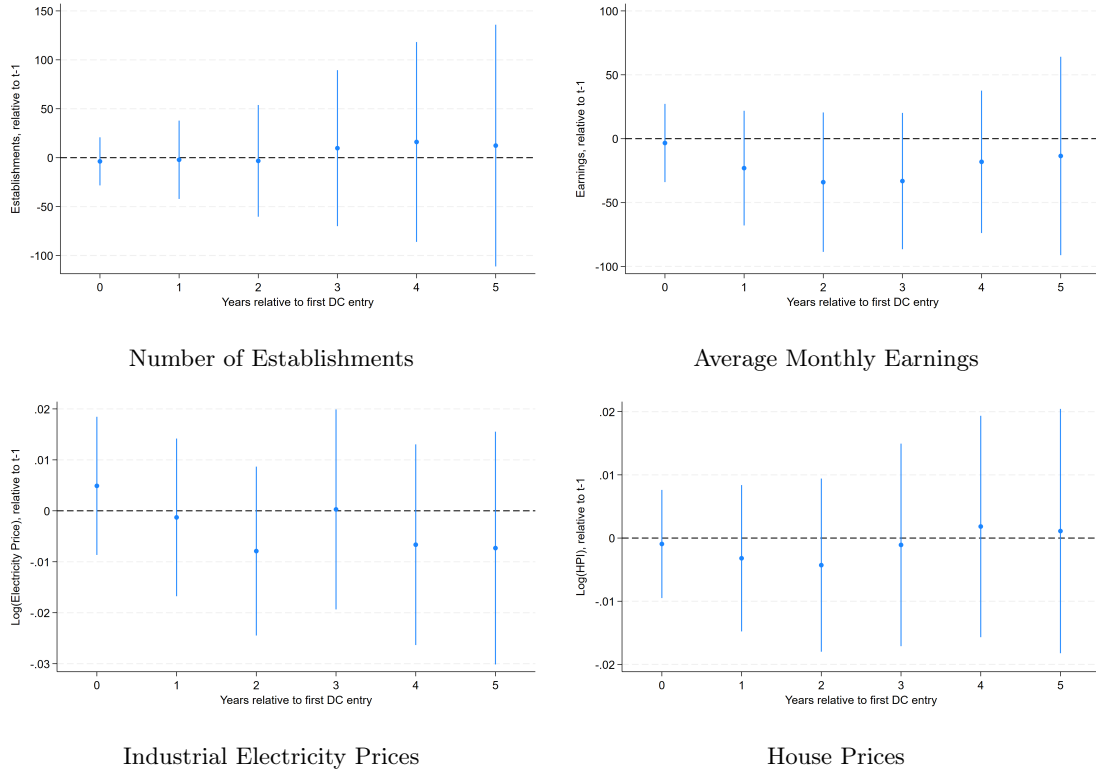
4.1.4 Business Activity and Earnings

We next examine local business activity, earnings, electricity prices, and house prices. Figure 6 reports effects on the number of non-construction establishments, average monthly earnings, industrial electricity prices, and house prices.

We find little evidence of changes in these outcomes. The number of establishments shows no clear response over the five years after activation. Average monthly earnings decline slightly, but the estimates are imprecise and statistically indistinguishable from zero.

We also find little evidence of changes in industrial electricity prices or house prices. The estimates remain close to zero throughout the post-activation period and are statistically insignificant.

Figure 6: **Effects of First Data Center Activation on Establishments, Earnings, Electricity Prices, and House Prices**



Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The donor pool consists of counties that never receive a data center during the sample period. The same SDiD unit and time weights estimated for non-construction employment are used for all outcomes. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Electricity prices and house prices are measured in logs. The earnings estimates use data from 2010 onward. Vertical bars report 95% jackknife confidence intervals.

Thus, the employment gains documented above do not coincide with detectable changes in establishments, average earnings, industrial electricity prices, or house prices. However, the baseline extensive-margin treatment captures the arrival of the first facility rather than the scale of installed capacity or subsequent expansion. We next exploit variation in installed capacity to examine how local outcomes vary with additional megawatts.

4.2 Intensive Margin: Installed Data Center Capacity

4.2.1 Specification

Data center investments vary enormously in scale. By 2023, among counties with at least one data center, the 25th percentile of installed capacity was 2.12 MW and the 75th percentile was 85.12 MW. Six percent of these counties host more than 500 MW. We now exploit this variation.

Our baseline sample here includes counties with and without data centers, excluding the top 5% by employment in 2000. The specification relates changes in local outcomes to differences in installed capacity across counties and over time, after absorbing persistent county differences, common state-year shocks, and observable determinants of economic growth or data center location. We allow local employment to adjust gradually by estimating horizon-specific local projections following [Jordà \(2005\)](#). For each horizon h , we estimate

$$y_{c,t+h} - y_{c,t-1} = \beta^h MW_{ct} + \delta_{st}^h + \omega_c^h + \mathbf{x}_{c,t}' \boldsymbol{\gamma}^h + \varepsilon_{ct}^h, \quad (5)$$

where $y_{c,t+h} - y_{c,t-1}$ is the cumulative change in outcome y from $t-1$ through horizon h . The variable MW_{ct} denotes installed data center capacity in county c and year t . County fixed effects, ω_c^h , absorb persistent differences across counties, while state-by-year fixed effects, δ_{st}^h , absorb common state-level shocks. The vector $\mathbf{x}_{c,t}$ includes baseline county characteristics interacted with linear year trends and is identical to those in Section 4.1.1. Standard errors are clustered by county.

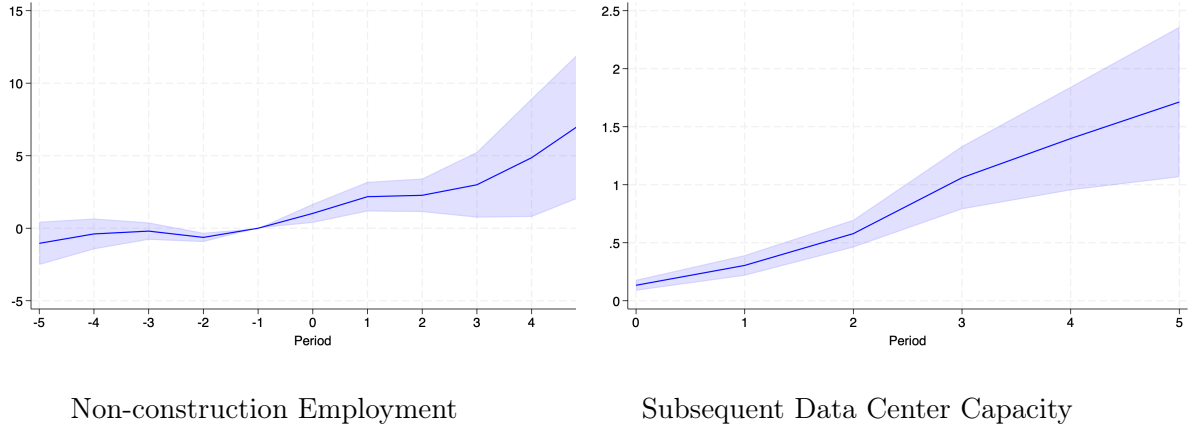
Our primary estimate for β^h is the cumulative change in local non-construction employment associated with one additional MW of installed capacity at horizon h . We estimate the model in levels rather than logarithms, as most county-year observations have zero installed data center capacity. Unlike the extensive-margin design, we do not interpret differences in installed capacity as conditionally random. The estimates therefore measure how local outcomes vary with installed data center capacity rather than the causal effect of an exogenous increase in capacity.

4.2.2 Employment Responses per Megawatt

Figure 7 shows how local employment and data center capacity evolve following an additional MW of installed capacity. The employment response builds gradually. By the five-year horizon, one additional MW is associated with 7.4 additional non-construction jobs. This pattern reflects local firms and labor markets adjusting over time, rather than employment rising immediately when new capacity comes online.

Installed capacity also predicts further data center investment. One additional MW is associated with 1.7 MW of subsequent capacity between t and $t + 6$. The initial MW is therefore associated with 2.7 MW of total capacity over the adjustment period. This follow-on investment matters for interpreting the employment response. We therefore normalize the five-year employment response by total effective MW in place after that period: $\frac{7.4}{1+1.7} \approx 2.74$.

Figure 7: **Local Employment and Subsequent Data Center Capacity per MW**



Notes: The figure reports estimates of equation (5) for horizons $h = 0, \dots, 5$, together with pre-treatment estimates for horizons $h = -5, \dots, -1$. The treatment variable is installed data center capacity, measured in MW, in county c and year t . The left panel reports the cumulative change in non-construction employment from $t - 1$ through horizon h associated with one additional MW of installed capacity. The right panel reports the cumulative change in installed data center capacity over the same horizon associated with one additional MW at time t . Shaded bands report 95% confidence intervals based on standard errors clustered by county. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1.

Table 2 reports the five-year estimates and examines the influence of Loudoun County, Virginia, the largest data center market in our sample. Excluding Loudoun increases the employment coefficient from 7.4 to 9.2 jobs per MW, while subsequent capacity growth falls from 1.7 to 0.9 MW.

Employment per effective MW consequently rises from 2.74 to 4.83. The baseline employment result is therefore not driven by Loudoun County.

One concern is that these local employment gains could reflect displacement from nearby counties. Appendix Table A1 examines employment in directly adjacent counties. We find no evidence of employment losses in neighboring counties as data center capacity expands.

Our sample ends in 2023, before the recent wave of much larger AI-related projects entered operation. The estimates describe the historical relationship between data center capacity and local employment over our sample period. They do not need to scale proportionally to the much larger facilities now under development.

Table 2: **Employment and Subsequent Data Center Capacity per MW**

<i>Panel A: Cumulative change in non-construction employment</i>		
	(1) Baseline sample	(2) Excl. Loudoun
Installed DC capacity (MW), year t	7.434*** (2.626)	9.214** (4.248)
Employment per effective MW	2.74	4.83
Within R-squared	0.179	0.179
Observations	52,560	52,542
<i>Panel B: Cumulative change in installed data center capacity</i>		
	(1) Baseline sample	(2) Excl. Loudoun
Installed DC capacity (MW), year t	1.713*** (0.330)	0.909*** (0.298)
Within R-squared	0.692	0.209
Observations	52,560	52,542

Notes: The table reports estimates of equation (5) at the five-year horizon. Panel A reports the cumulative change in non-construction employment from $t-1$ through $t+5$ associated with one additional MW of installed data center capacity in year t . Panel B reports the change in installed data center capacity from t through $t+6$ associated with one additional MW of installed capacity in year t . The baseline sample excludes the top 5% of counties by employment in 2000. Column (2) additionally excludes Loudoun County, Virginia. “Employment per effective MW” equals the coefficient in Panel A divided by one plus the corresponding coefficient in Panel B and is reported as an accounting normalization without a joint standard error. All specifications include county and state-by-year fixed effects, Republican vote share, and baseline county characteristics interacted with linear year trends: industrial electricity prices, internet infrastructure, 1980–2000 employment growth, the share with a bachelor’s degree, the foreign-born share, the share under age 25, the share of land covered by water, January temperature, and scenic amenity. Standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2.3 Incumbent Expansion and New-Operator Entry

Installed capacity predicts subsequent data center investment. Does this relationship reflect continued investment by incumbent operators or new-operator entry?

Table 3: **Employment and Subsequent Data Center Capacity by Operator Type**

Panel A: Cumulative change in non-construction employment, by operator type of MW added

	(1) Jobs ($t - 1$ to $t + 5$)
MW from incumbent operator (same-operator expansion)	3.910 (4.128)
MW from new operator (first entry to county)	19.33 (15.31)
Within R-squared	0.179
Observations	52,560

Panel B: Cumulative change in installed data center capacity, by operator type of MW added

	(1) MW in place (t to $t + 6$)
MW from incumbent operator (same-operator expansion)	2.074*** (0.427)
MW from new operator (first entry to county)	0.798 (0.583)
Within R-squared	0.704
Observations	52,560

Panel C: Operator decomposition of subsequent data center capacity

	(1) Existing-operator MW (t to $t + 6$)	(2) New-operator MW (t to $t + 6$)
MW from incumbent operator (same-operator expansion)	1.453*** (0.254)	0.620*** (0.190)
MW from new operator (first entry to county)	1.583*** (0.527)	-0.777*** (0.143)
Within R-squared	0.758	0.329
Observations	52,560	52,560

Notes: Panels A and B split installed data center capacity into mutually exclusive categories based on operator identity. Incumbent-operator capacity is capacity added by an operator already present in the county in an earlier year. New-operator capacity is capacity added by an operator in its first year of presence in the county. Capacity associated with operators reported by Aterio as “Company Not Disclosed” enters separately in the regressions but is not reported because individual operators cannot be identified. Panel A reports the cumulative change in non-construction employment from $t - 1$ to $t + 5$. Panel B reports the cumulative change in total installed data center capacity from t to $t + 6$. Panel C decomposes the Panel B outcome according to whether subsequent capacity is added by an operator already present in the county as of year t (existing-operator MW) or by an operator not yet present as of year t (new-operator MW). An undisclosed-operator category completes the partition but is not reported. All specifications use the baseline sample and include county and state-by-year fixed effects and the controls described in Section 4.2.1. Standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Follow-on capacity can arise from an incumbent operator expanding an existing campus in phases. Alternatively, existing data center activity may attract new operators nearby. New-operator entry provides stronger evidence of colocation forces that extend beyond the incumbent operator (Ellison et al., 2010).

To distinguish between these channels, we classify new capacity as incumbent expansion when the operator was already present in the county, or new-operator entry otherwise. We then enter both measures simultaneously in a local projection specification.

Panel A of Table 3 reports the employment results. One MW associated with new-operator entry predicts 19.3 additional non-construction jobs, compared with 3.9 jobs for incumbent expansion. Both estimates are imprecise, however, and we cannot reject equality between the two coefficients.

Panel B examines subsequent data center investment. One additional MW added by an incumbent operator is associated with 2.1 MW of subsequent total county capacity. In contrast, one MW added by a new operator predicts 0.8 MW of subsequent capacity, and the estimate is not statistically significant.

Panel C decomposes subsequent capacity according to whether it is added by an operator already present in the county as of year t or by an operator not yet present. Of the 2.1 MW of subsequent capacity associated with one MW of incumbent investment, 1.45 MW comes from operators already present and 0.62 MW from operators not yet present. Both estimates are statistically significant. Therefore, most follow-on investment reflects continued investment by existing operators, but incumbent capacity also predicts subsequent investment by new operators.

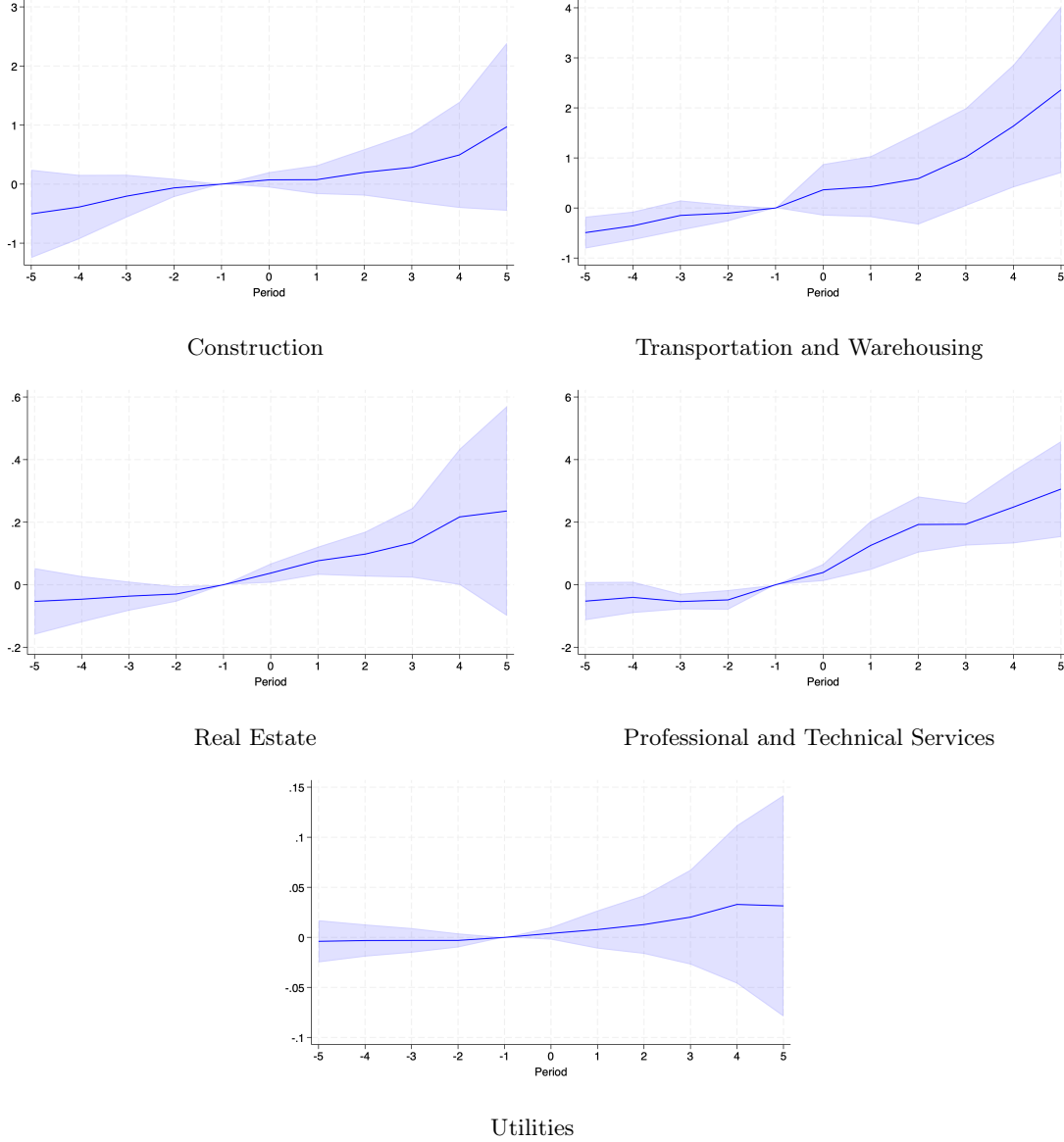
The pattern following new-operator entry is different. One MW of new-operator capacity predicts 1.58 MW of subsequent capacity from operators present as of year t , but -0.78 MW from other operators. The first relationship partly reflects the classification. Once an operator enters a county, its subsequent expansion is classified as capacity from an operator already present. The negative coefficient for other operators provides little evidence that new-operator entry is followed by successive entry from additional operators. Overall, the results point to substantial persistence in data center investment, together with some evidence that established data center activity predicts investment by new operators.

4.2.4 Sectoral Employment Effects

Figure 8 shows the sectors with the strongest positive employment responses. Employment rises most in transportation and warehousing and in professional, scientific, and technical services. Real

estate employment also shows a positive but smaller response. These sectors provide logistics, property, technical, and other services that can complement data center investment, consistent with local production linkages.

Figure 8: **Employment Effects per MW by Selected Sector**

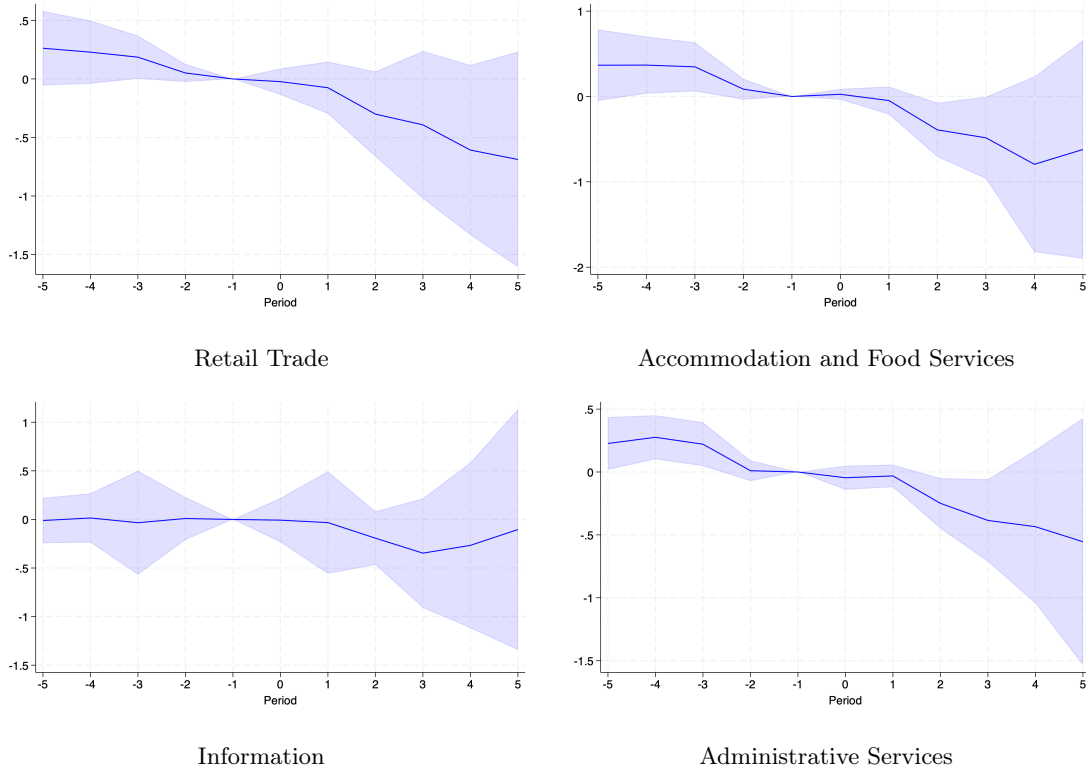


Notes: The figure reports estimates of equation (5) for selected sectors. For all sectors except construction, each panel shows the cumulative change in sector employment from $t-1$ through horizon h associated with one additional MW of installed data center capacity in county c and year t . For construction, employment is shifted back by 3 years relative to installed capacity to capture activity associated with building the facility before the capacity becomes operational. Shaded bands report 95% confidence intervals based on standard errors clustered by county. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1.

We again shift construction employment back three years relative to installed capacity to account for the construction period. The estimates are positive but imprecise. Utilities also show a small positive response, but the estimates are not statistically different from zero.

Figure 9 shows the employment responses in retail, accommodation and food services, and administrative services. The estimates are generally negative but statistically insignificant. We therefore find little evidence of broad employment growth through a local consumption channel. The negative point estimates in accommodation and food services and retail are largely driven by Loudoun County.

Figure 9: **Employment Effects per MW in Other Local Sectors**



Notes: The figure reports estimates of equation (5) for selected sectors. Each panel shows the cumulative change in sector employment from $t-1$ through horizon h associated with one additional MW of installed data center capacity in county c and year t . Shaded bands report 95% confidence intervals based on standard errors clustered by county. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1.

Table 4 reports the estimates at horizon $h = 5$ excluding Loudoun County. In the baseline sample, one additional MW is associated with 2.36 additional transportation and warehousing jobs and 3.06 additional professional and technical services jobs. Both estimates are statistically

significant. Excluding Loudoun yields corresponding estimates of 3.09 and 1.46 jobs per MW. Real estate also becomes positive and statistically significant outside Loudoun, at 0.56 jobs per MW.

Table 4: **Employment Effects per MW by Sector**

<i>Panel A: Baseline sample</i>									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Constr.	Transp./Wareh.	Real Estate	Prof./Tech.	Utilities	Retail	Accom./Food	Information	Admin.
Installed DC capacity (MW), year t	1.809 (1.666)	2.359*** (0.846)	0.236 (0.171)	3.057*** (0.781)	0.0315 (0.0564)	-0.688 (0.469)	-0.621 (0.652)	-0.104 (0.632)	-0.554 (0.501)
Observations	49,640	52,560	52,560	52,560	52,560	52,560	52,560	52,560	52,560
<i>Panel B: Baseline sample excluding Loudoun County</i>									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Constr.	Transp./Wareh.	Real Estate	Prof./Tech.	Utilities	Retail	Accom./Food	Information	Admin.
Installed DC capacity (MW), year t	2.541 (2.238)	3.090** (1.472)	0.563*** (0.213)	1.459** (0.641)	0.0570 (0.101)	-0.177 (0.626)	0.491 (0.323)	-0.730 (0.774)	-0.528 (0.800)
Observations	49,623	52,542	52,542	52,542	52,542	52,542	52,542	52,542	52,542

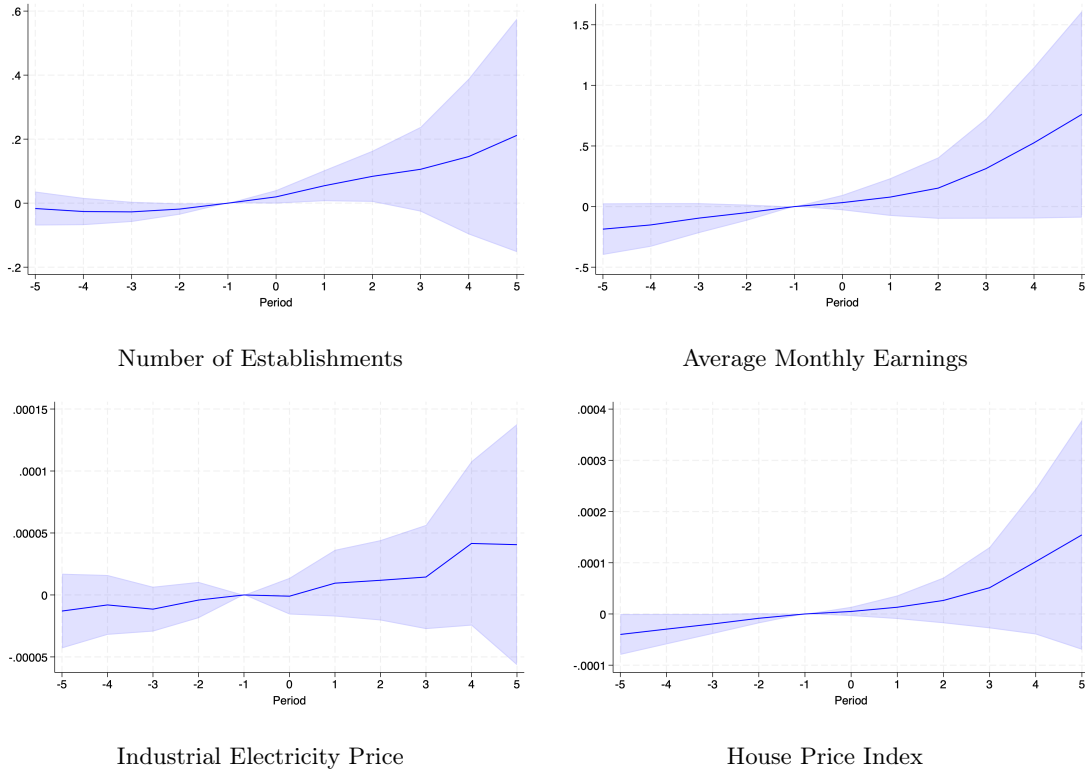
Notes: The table reports estimates of equation (5) for employment by sector. For all sectors except construction, the dependent variable is the cumulative change in employment from $t-1$ to $t+5$. For construction, employment is shifted back by 3 years relative to installed capacity to capture activity associated with building the facility before the capacity becomes operational. Panel A uses the baseline sample, which excludes the top 5% of counties by employment in 2000. Panel B additionally excludes Loudoun County, Virginia. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1. Standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The remaining sectors show no statistically significant employment response in either sample. The intensive-margin results need not mirror the activation results in Section 4.1.3. Activation captures data center entry, while the estimates here capture differences in the scale of installed capacity. Overall, the employment response is concentrated in a relatively narrow set of industries linked to data center activity. We find much less evidence that the investment spreads broadly into local consumer services. These eight sectors, however, do not exhaust total non-construction employment. Summing their point estimates in the baseline sample yields approximately 3.7 jobs per MW, about half of the 7.4 jobs per MW aggregate coefficient reported in Table 2. The remaining non-construction sectors, not reported individually in Table 4, account for the balance.

4.2.5 Business Activity, Earnings, and Local Prices

Figure 10 examines business activity, earnings, electricity prices, and house prices. Unlike the extensive-margin estimates, establishment counts increase with installed data center capacity. By the five-year horizon, one additional MW is associated with approximately 0.2 additional establishments, although the estimate is not statistically significant. Normalizing by the 2.7 MW of effective capacity gives about 0.07 additional establishments per effective MW.

Figure 10: **Effects per MW on Establishments, Earnings, Electricity Prices, and House Prices**



Notes: The figure reports estimates of equation (5) for additional local economic outcomes. Each panel shows the cumulative change in the indicated outcome from $t - 1$ through horizon h associated with one additional MW of installed data center capacity in county c and year t . Electricity prices and the house price index enter in logs. Shaded bands report 95% confidence intervals based on standard errors clustered by county. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1.

Average monthly earnings also increase with installed capacity, although the baseline estimate is only marginally significant. At the five-year horizon, one additional MW is associated with approximately \$0.76 higher monthly earnings per worker. Applying the same accounting normalization gives approximately \$3.4 in annual earnings per worker for each effective MW.¹⁹ We find little evidence that installed data center capacity is associated with changes in county-level industrial electricity prices.

House prices show a similar pattern in the baseline sample. The point estimates become more positive at longer horizons but remain imprecisely estimated. At the five-year horizon, one additional MW is associated with a 0.015% increase in house prices, but the estimate is not statistically

¹⁹The calculation multiplies the five-year monthly earnings coefficient of \$0.76 by 12 and divides by 2.71 MW of effective capacity.

significant.

Table 5 reports the five-year estimates for all four outcomes and examines their sensitivity to excluding Loudoun County. The establishment coefficient increases from 0.21 to 0.38 establishments per MW but remains statistically insignificant. The earnings coefficient increases from \$0.76 to \$1.63 per month per MW and becomes statistically significant at the 1% level. The house-price estimate increases from 0.015% to 0.038% per MW and becomes statistically significant at the 5% level. The electricity-price estimate remains small and statistically insignificant. The house-price magnitude is consistent with [Priest \(2026\)](#), who finds economically small, slightly positive effects of data center proximity on house prices in Virginia using property-level data.

Table 5: **Business Activity, Earnings, and Local Prices per MW**

<i>Panel A: Baseline sample</i>				
	(1)	(2)	(3)	(4)
	Non-constr. establishments	QWI earnings	Industrial electricity price	Log HPI
Installed DC capacity (MW)	0.212 (0.186)	0.761* (0.434)	0.00004 (0.00005)	0.000154 (0.000114)
Observations	52,560	51,653	52,560	45,081
<i>Panel B: Excluding Loudoun County</i>				
	(1)	(2)	(3)	(4)
	Non-constr. establishments	QWI earnings	Industrial electricity price	Log HPI
Installed DC capacity (MW)	0.378 (0.356)	1.626*** (0.473)	0.00001 (0.00007)	0.000378** (0.000153)
Observations	52,542	51,635	52,542	45,063

Notes: The table reports estimates of equation (5) at the five-year horizon for county-level business activity, earnings, electricity prices, and house prices. The dependent variable is the cumulative change in each outcome from $t - 1$ to $t + 5$. Panel A reports the baseline sample, which excludes the top 5% of counties by employment in 2000. Panel B additionally excludes Loudoun County, Virginia. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1. Standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.3 Robustness

We report additional extensive-margin results in Appendix A.4. Appendix A.4.1 extends the sectoral analysis to additional industries. Appendix A.4.2 re-estimates the employment results using the Census Quarterly Workforce Indicators (QWI). The QWI estimates support the increase in non-construction employment and show positive point estimates in several sectors linked to data center activity, although the sectoral estimates are less precise than in CBP.

Appendix A.4.3 re-estimates the response to first data center activation using the local projection difference-in-differences estimator of [Dube et al. \(2025\)](#). Unlike SDiD, LP-DiD compares newly

treated counties with counties that have not yet been treated and relies on conditional parallel trends rather than reweighting control counties to match treated counties’ pre-treatment outcome paths. Some sectors, including transportation and warehousing, retail, and accommodation and food services, show evidence of differential pre-treatment trends. The post-treatment estimates are also generally larger than under SDiD. These patterns are consistent with treated counties following steeper trajectories before data center entry. SDiD addresses this concern by reweighting control counties to more closely match treated counties’ pre-treatment outcome paths. We therefore prefer SDiD for the extensive-margin analysis.

We report additional intensive-margin tests in Appendix A.5. Appendix A.5.1 examines employment in adjacent counties. We find no evidence that employment gains in host counties come at the expense of neighboring counties. Appendix A.5.2 sequentially excludes the largest data center hubs. The employment response remains positive, while the positive follow-on capacity response weakens and turns negative as we remove the largest hubs. Follow-on capacity growth is therefore concentrated in major data center markets.

Finally, Appendix A.5.3 re-estimates the intensive-margin employment results using QWI. Employment increases in transportation and warehousing, professional and technical services, and total non-construction employment. We continue to find little evidence of broad employment growth across local consumer-service sectors.

5 Conclusion

Data centers are among the most capital-intensive investments in the modern economy, yet they employ relatively few workers directly. This raises a simple question: do these investments generate jobs elsewhere in the local economy?

Our results indicate that they do. Counties gain employment after their first data center opens, and employment continues to rise with installed computing capacity. At the five-year horizon, one additional MW is associated with 7.4 additional non-construction jobs; because the initial MW predicts another 1.7 MW of capacity, this corresponds to 2.74 jobs per effective MW.

Employment grows most in transportation and warehousing, professional and technical services, and real estate. By contrast, we find little employment growth in retail, restaurants, and other consumer-service sectors. This pattern is consistent with production linkages to complementary local firms playing a larger role than a broad increase in local consumer demand.

We find weaker evidence for a second potential mechanism, colocation. Installed capacity predicts substantial subsequent capacity growth in the full sample, but this relationship is concentrated around investment by incumbent operators and disappears as we exclude the largest hubs. Existing capacity therefore predicts continued investment by incumbent operators, but our results provide weaker evidence that it systematically attracts new operators.

Effects beyond employment are less precise. Average earnings rise modestly, particularly when we exclude Loudoun County. We find little evidence of changes in industrial electricity prices, and house prices increase only mildly in specifications that exclude Loudoun County. Importantly, these estimates do not capture the full local costs of data center development, including demands on electricity infrastructure, water, land, and other public resources.

Our results describe an industry that is changing rapidly. The sample ends in 2023, before the latest wave of AI-related investment. New facilities are much larger, and direct employment per MW appears to decline with scale. The local employment effects we estimate need not increase in proportion to the size of these new projects—this distinction matters for policy. The relevant question is not whether a data center employs many workers inside the facility. It is whether the investment creates enough economic activity around it to justify the resources used to attract and accommodate it. We find meaningful local employment gains concentrated in industries linked to data center activity and these gains can serve as one input to the cost-benefit calculations around data centers. However, we do not estimate whether these gains justify public incentives and infrastructure costs or how they compare with environmental externalities.

References

- Abadie, A., 2021. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature* 59, 391–425. doi:[10.1257/jel.20191450](https://doi.org/10.1257/jel.20191450).
- Acemoglu, D., Restrepo, P., 2018. The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review* 108, 1488–1542. doi:[10.1257/aer.20160696](https://doi.org/10.1257/aer.20160696).
- Almagro, M., Domínguez-Iino, T., 2025. Location Sorting and Endogenous Amenities: Evidence from Amsterdam. *Econometrica* 93, 1031–1071. doi:[10.3982/ECTA21394](https://doi.org/10.3982/ECTA21394).
- Alvarez, F.E., Argente, D., Chow, J., 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. Cambridge, MA. URL: <https://www.nber.org/papers/w35194>, doi:[10.3386/w35194](https://doi.org/10.3386/w35194).
- Amior, M., 2025. The contribution of immigration to local labor market adjustment. *Journal of Labor Economics* 43, 1169–1206.
- Arkhangelsky, D., Athey, S., Hirshberg, D.A., Imbens, G.W., Wager, S., 2021. Synthetic Difference-in-Differences. *American Economic Review* 111, 4088–4118. doi:[10.1257/aer.20190159](https://doi.org/10.1257/aer.20190159).
- Bahar, D., Wright, G.C., 2026. New Evidence on Data Center Employment Effects. Brookings Institution. URL: <https://www.brookings.edu/articles/new-evidence-on-data-center-employment-effects/>.
- Baum-Snow, N., Gendron-Carrier, N., Pavan, R., 2024. Local Productivity Spillovers. *American Economic Review* 114, 1030–1069. doi:[10.1257/aer.20211198](https://doi.org/10.1257/aer.20211198).
- Black, D., Henderson, V., 1999. A theory of urban growth. *Journal of political economy* 107, 252–284.
- Brandsaas, E.E., Garcia, D., Kurtzman, R.J., Nichols, J.B., Zyttek, A., 2025. Estimating Aggregate Data Center Investment with Project-Level Data. Finance and Economics Discussion Series 2025-109. Board of Governors of the Federal Reserve System. Washington, DC. doi:[10.17016/FEDS.2025.109](https://doi.org/10.17016/FEDS.2025.109).
- Büchler, S.C., Niu, D., Thompson, A.K., Zheng, S., 2024. The Impact of Human Capital and Housing Supply on Urban Growth. *Urban Studies* 61, 214–230. doi:[10.1177/00420980231159687](https://doi.org/10.1177/00420980231159687).
- Burchfield, M., Overman, H.G., Puga, D., Turner, M.A., 2006. Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics* 121, 587–633.
- CAIDA, 2026. CAIDA Internet Exchange Points (IXPs) Dataset. URL: <https://www.caida.org/catalog/datasets/ixps/>.
- Carlino, G.A., Saiz, A., 2019. Beautiful City: Leisure Amenities and Urban Growth. *Journal of Regional Science* 59, 369–408. doi:[10.1111/jors.12438](https://doi.org/10.1111/jors.12438).
- Carlo, A., Rolheiser, L., 2026. Data Centred: The Shifting Landscape of Canada’s Digital Infrastructure. URL: <https://papers.ssrn.com/abstract=6464099>, doi:[10.2139/ssrn.6464099](https://doi.org/10.2139/ssrn.6464099). working paper.

- Chen, J., Roth, J., 2024. Logs with Zeros? Some Problems and Solutions. *Quarterly Journal of Economics* 139, 891–936. doi:[10.1093/qje/qjad054](https://doi.org/10.1093/qje/qjad054).
- Ciccia, D., 2024. A Short Note on Event-Study Synthetic Difference-in-Differences Estimators. URL: <https://arxiv.org/abs/2407.09565>. arXiv:2407.09565.
- Du, R., Sun, X.A., Zheng, S., 2026. Bits and Mortar: The Microspatial Economics of Data Centers. URL: <https://papers.ssrn.com/abstract=6863998>, doi:[10.2139/ssrn.6863998](https://doi.org/10.2139/ssrn.6863998). working paper.
- Dube, A., Girardi, D., Jordà, Ò., Taylor, A.M., 2025. A Local Projections Approach to Difference-in-Differences Event Studies. *Journal of Applied Econometrics* 40, 741–758. doi:[10.1002/jae.70000](https://doi.org/10.1002/jae.70000).
- Duranton, G., Puga, D., 2004. Micro-Foundations of Urban Agglomeration Economies, in: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics*. Elsevier. volume 4, pp. 2063–2117. doi:[10.1016/S1574-0080\(04\)80005-1](https://doi.org/10.1016/S1574-0080(04)80005-1).
- Eckert, F., Fort, T.C., Schott, P.K., Yang, N.J., 2020. Imputing Missing Values in the US Census Bureau’s County Business Patterns. Working Paper 26632. National Bureau of Economic Research. Cambridge, MA. doi:[10.3386/w26632](https://doi.org/10.3386/w26632).
- Ellison, G., Glaeser, E.L., Kerr, W.R., 2010. What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review* 100, 1195–1213. doi:[10.1257/aer.100.3.1195](https://doi.org/10.1257/aer.100.3.1195).
- Faverio, M., Kikuchi, E., 2026. What the Data Says about Americans’ Views of Artificial Intelligence. Pew Research Center. URL: <https://www.pewresearch.org/short-reads/2026/03/12/what-the-data-says-about-americans-views-of-artificial-intelligence/>.
- Federal Housing Finance Agency, 2026. FHFA House Price Index. URL: <https://www.fhfa.gov/data/hpi>.
- Gargano, A., Giacoletti, M., 2025. Subsidizing Digital Infrastructure: Local Economic Effects of Data Center Investments. URL: <https://papers.ssrn.com/abstract=5881105>, doi:[10.2139/ssrn.5881105](https://doi.org/10.2139/ssrn.5881105). working paper.
- Glaeser, E.L., Kolko, J., Saiz, A., 2001. Consumer city. *Journal of Economic Geography* 1, 27–50.
- Glaeser, E.L., Saiz, A., 2004. The Rise of the Skilled City. *Brookings-Wharton Papers on Urban Affairs* 2004, 47–105. doi:[10.1353/urb.2004.0005](https://doi.org/10.1353/urb.2004.0005).
- Glaeser, E.L., Scheinkman, J., Shleifer, A., 1995. Economic growth in a cross-section of cities. *Journal of monetary economics* 36, 117–143.
- Gour, N., Ortiz, L., Maibach, E., 2026. Health and Environmental Risks from Data Centers: Implications for Communication and Community Engagement. *Frontiers in Climate* 8, 1648912. doi:[10.3389/fclim.2026.1648912](https://doi.org/10.3389/fclim.2026.1648912).
- Greenstone, M., Hornbeck, R., Moretti, E., 2010. Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy* 118, 536–598. doi:[10.1086/653714](https://doi.org/10.1086/653714).
- Hicks, M.J., 2025. Data Centers and Local Job Creation. Substack. URL: <https://michaelhicks.substack.com/p/data-centers-and-local-job-creation>.

- International Energy Agency, 2025. Data Centres and Networks. URL: <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>.
- Jordà, Ò., 2005. Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review* 95, 161–182. doi:[10.1257/0002828053828518](https://doi.org/10.1257/0002828053828518).
- Kay, O., Reaser, R., Taylor, R., 2026. Processing Power: The Local Economic Impact of Data Centers. Working Paper 2606. Federal Reserve Bank of Dallas. doi:[10.24149/wp2606](https://doi.org/10.24149/wp2606).
- Kline, P., Moretti, E., 2014. Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics* 129, 275–331. doi:[10.1093/qje/qjt034](https://doi.org/10.1093/qje/qjt034).
- Marshall, A., 1920. *Principles of Economics*. 8 ed., Macmillan and Co., London. Book IV, chap. X, Industrial Organization.
- Mathur, A., Sharma, S., 2026. How Data Centers Create Noise Pollution. NPR. URL: <https://www.npr.org/2026/04/20/nx-s1-5545074/data-center-noise-pollution>.
- MIT Election Data and Science Lab, 2026. MIT Election Data and Science Lab. URL: <https://electionlab.mit.edu/>.
- Moretti, E., 2004. Human Capital Externalities in Cities, in: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics*. Elsevier. volume 4, pp. 2243–2291. doi:[10.1016/S1574-0080\(04\)80008-7](https://doi.org/10.1016/S1574-0080(04)80008-7).
- Mytton, D., 2021. Data Centre Water Consumption. *npj Clean Water* 4, 11. doi:[10.1038/s41545-021-00101-w](https://doi.org/10.1038/s41545-021-00101-w).
- National Conference of State Legislatures, 2026. Subsidizing Data Centers: State Approaches. URL: <https://www.ncsl.org/technology-and-communication/subsidizing-data-centers-state-approaches>.
- O’Leary, S., 2025. Why Data Centers Are a Bad Deal for Local Communities. Ohio River Valley Institute. URL: <https://ohiorivervalleyinstitute.org/why-data-centers-are-a-bad-deal-for-local-communities/>.
- Priest, A., 2026. Not in My Back Yard! The Effects of Data Centers on Housing Prices. URL: <https://papers.ssrn.com/abstract=6314620>, doi:[10.2139/ssrn.6314620](https://doi.org/10.2139/ssrn.6314620). working paper.
- Rosenthal, S.S., Strange, W.C., 2004. Evidence on the Nature and Sources of Agglomeration Economies, in: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics*. Elsevier. volume 4, pp. 2119–2171. doi:[10.1016/S1574-0080\(04\)80006-3](https://doi.org/10.1016/S1574-0080(04)80006-3).
- Ryu, A., Hiatt, S.R., 2025. Data Center Employment Forecast Analysis. Technical Report. Hamm Institute for American Energy, Oklahoma State University. URL: https://hamminstitute.org/site-files/documents/data_center_workforce.pdf.
- Saiz, A., Salazar, A., Bernard, J., 2018. Crowdsourcing Architectural Beauty: Online Photo Frequency Predicts Building Aesthetic Ratings. *PLOS ONE* 13, e0194369. doi:[10.1371/journal.pone.0194369](https://doi.org/10.1371/journal.pone.0194369).

- Sarte, K.A., Gribbin, M., Miller, E., Berday-Sacks, S., Hopkins, K., Saunders, S., 2024. Data Centers in Virginia. JLARC Report 598. Joint Legislative Audit and Review Commission, Commonwealth of Virginia. URL: <https://jlarc.virginia.gov/pdfs/reports/Rpt598-2.pdf>.
- Shehabi, A., Smith, S.J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M.A.B., Holecek, B., Koomey, J., Masanet, E., Sartor, D., 2024. 2024 United States Data Center Energy Usage Report. Technical Report LBNL-2001637. Lawrence Berkeley National Laboratory. doi:[10.71468/P1WC7Q](https://doi.org/10.71468/P1WC7Q).
- U.S. Bureau of Economic Analysis, 2026. Personal Income by County. URL: <https://www.bea.gov/data/income-saving/personal-income-by-county>.
- U.S. Census Bureau, 2026. Quarterly Workforce Indicators. URL: <https://www.census.gov/data/datasets/time-series/qwi.html>.
- U.S. Department of Agriculture, Economic Research Service, 2025. Natural Amenities Scale. URL: <https://www.ers.usda.gov/data-products/natural-amenities-scale>.
- U.S. Energy Information Administration, 2026. Electricity Data. URL: <https://www.eia.gov/electricity/data.php>.
- White House, 2025a. Accelerating Federal Permitting of Data Center Infrastructure. URL: <https://www.whitehouse.gov/presidential-actions/2025/07/accelerating-federal-permitting-of-data-center-infrastructure/>.
- White House, 2025b. Ensuring American Leadership in Artificial Intelligence. URL: <https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/>.

A Appendix

A.1 Empirical Challenges

Estimating the local economic effects of data centers raises several empirical challenges. Data centers employ few workers directly, differ enormously in scale, and locate selectively. Construction precedes operation, while jobs associated with the investment can appear outside the facility and several years after it opens. These features affect both the outcome we measure and the comparison needed to interpret it.

Crowding out and spillovers. Jobs reported by data center operators need not measure net local job creation. Consider an electrician who previously worked as an independent contractor in the same county and is subsequently hired by a data center. The facility reports an additional employee, but county employment has not increased. Direct hiring can therefore displace employment elsewhere in the local labor market. The opposite problem arises with spillovers. A data center may purchase services from a local firm, creating jobs that never appear on the operator’s payroll. We therefore measure county-wide employment and examine where those jobs appear within the local economy.

Construction employment. Construction can account for many jobs associated with a data center, but these jobs occur before the facility opens and are temporary. Combining construction and operating employment would mix two different responses. Our main outcome therefore excludes construction. We estimate construction employment separately and shift its timing back three years relative to activation or installed capacity.

Logs versus levels. The choice between logs and levels is consequential in this setting. Consider the same 10 MW data center locating in a county with 10,000 workers and another with one million. There is little reason to expect the facility to create one hundred times as many jobs in the larger county. A proportional specification nevertheless moves in this direction. Since we observe the physical scale of the investment in MW, we instead ask how many additional jobs are associated with it.

Levels have their own drawback. Normal employment changes are much larger in large counties, and county employment in the United States is highly skewed. A change of 500 jobs may be substantial in one county and indistinguishable from normal fluctuations in another. This motivates

our exclusion of the largest 5% of counties and, in the extensive-margin analysis, the use of synthetic controls with similar pre-treatment employment paths. Clustering standard errors by county allows for heteroskedasticity and serial correlation within counties.

Entry versus capacity. A county’s first data center opening provides a clear treatment date. What happens afterward is less discrete. Existing facilities expand, new operators enter, and these investments can range from a few MW to hundreds of MW. Treating each change as an equivalent event would ignore economically important differences in scale. We therefore use first activation for the extensive-margin analysis and installed MW for the intensive margin. Most county-year observations have zero installed capacity, so using MW in levels also avoids the arbitrary treatment of zeros required by a logarithmic specification.

Timing. Employment need not respond in the year a data center becomes operational. Our extensive-margin estimates illustrate the problem directly. Non-construction employment changes little during the first two years and becomes larger only after three years. A specification focused only on the year of activation would miss this pattern. We therefore trace employment over several years in both the event-study and local-projection specifications.

Selection and predictable growth. Data centers do not locate randomly. Power infrastructure, electricity costs, fiber connectivity, land availability, tax incentives, and other local conditions shape where facilities are built. Several of these characteristics also predict future economic growth. A county that receives a data center may therefore have evolved differently even without the investment.

Some of these differences are predictable from initial conditions. The urban economics literature shows that human capital predicts subsequent local growth (Moretti, 2004; Glaeser and Saiz, 2004; Büchler et al., 2024). Immigration (Amior, 2025), climate (Glaeser et al., 1995), amenities (Glaeser et al., 2001; Carlini and Saiz, 2019), access to water (Burchfield et al., 2006), and local policies (Black and Henderson, 1999) can also shape local growth. Moreover, county growth is persistent, so past employment growth contains information about future growth (Glaeser et al., 1995). Our specifications therefore allow baseline county characteristics, measured in or prior to 2000, to generate differential linear trends. These characteristics include past employment growth and variables that predict both data center location and subsequent local economic growth.

For first activation, synthetic difference-in-differences further addresses predictable differences

in growth by placing greater weight on never-treated counties with employment paths similar to treated counties before activation. This lets the comparison group resemble treated counties not only in observed baseline characteristics but also in pre-treatment employment dynamics.

Time-varying local shocks. Baseline characteristics and pre-treatment employment paths cannot capture every shock around data center investment. A county receiving a large capacity addition may simultaneously experience other investments or changes in economic conditions. State-by-year fixed effects absorb shocks common to counties within the same state in a given year, while county fixed effects absorb persistent differences across counties.

The local-projection specification measures cumulative changes in outcomes from $t - 1$ through each horizon. This removes differences in outcome levels that predate the capacity observed at t , but it does not eliminate time-varying county-specific shocks. In particular, installed capacity may correlate with other local investments or economic changes over the same period. Unlike the extensive-margin design, we therefore do not interpret differences in installed capacity as conditionally random. The intensive-margin estimates measure how local outcomes vary with installed data center capacity rather than the causal effect of an exogenous increase in capacity.

Measurement error. Both sides of the analysis are measured with error. For data centers, errors in activation dates affect treatment timing, while errors in installed MW affect measured treatment intensity. This is particularly relevant for facilities built in phases, where facility opening and capacity additions need not occur at the same time. Our asset-level data allow us to distinguish activation from subsequent capacity additions. For 153 facilities with missing activation years in the original data, student research assistants identified activation years through manual research.

Employment creates a different measurement problem. County Business Patterns suppresses some county-industry employment cells, making measurement particularly difficult for narrow industries and smaller counties. We use the imputations for missing cells provided by [Eckert et al. \(2020\)](#), which combine reported employment ranges with adding-up constraints across industries and geographic units. We also repeat the main employment specifications using the Census Quarterly Workforce Indicators (QWI). The similar patterns across the two sources reduce the likelihood that CBP-specific measurement error drives our results.

A.2 A Model of Local Economic Effects

We develop the microeconomic foundations for the conceptual framework in Section 2. Data centers combine large capital investments with little direct labor. Their broader local effects operate through demand for inputs and services and through spending by workers and residents. Local markets determine how these forces translate into employment, wages, and local costs.

A.2.1 Environment and Production Linkages

Consider a local economy with three sectors. Data centers, denoted by D , operate computing capacity K and employ local labor L_D . A linked sector, S , supplies inputs and services used by data centers. These include maintenance, security, networking, real estate, and professional and technical services. A nontradable sector, N , supplies goods and services consumed locally. We take installed data center capacity K as the source of the local demand shock.

Data centers require relatively little labor to operate their installed capacity. We write direct employment as

$$L_D = \lambda(K)K, \quad \lambda'(K) < 0, \quad (6)$$

where $\lambda(K)$ measures direct employment per unit of installed capacity. A small $\lambda(K)$ captures the labor-light nature of data centers. Provider-level evidence suggests that employment intensity also declines with facility scale.²⁰ Larger facilities can therefore add substantial computing capacity without increasing direct employment proportionally.

Data centers also purchase inputs and services from local firms. We capture this production linkage through linked-sector labor demand:

$$L_S^d = A_S K^\theta w_S^{-\eta_S}, \quad \theta > 0, \quad \eta_S > 0, \quad (7)$$

where w_S denotes the linked-sector wage. The parameter A_S shifts the level of local labor demand, while θ measures how strongly that demand responds to installed capacity. The parameter η_S governs the response of labor demand to wages.

Installed computing capacity can therefore increase local labor demand even when data centers themselves employ few workers. Direct employment operates through L_D , while local production linkages shift labor demand in S . The more inputs and services data centers purchase locally, the

²⁰See [Ryu and Hiatt \(2025\)](#).

larger the potential employment response outside the facility.

A.2.2 Local Expenditure and Equilibrium

The initial employment response can spread through local spending. Workers in data centers and linked firms earn income, some of which they spend on local goods and services. Let $B_j = w_j L_j$ denote the wage bill in sector j . Households spend a fraction α of local labor income on nontradables. Labor therefore earns

$$B_N = \alpha \beta_N (B_D + B_S + B_N), \quad (8)$$

where β_N is labor's share of revenue in the nontradable sector. Solving for B_N gives

$$B_N = \mu (B_D + B_S), \quad \mu \equiv \frac{\alpha \beta_N}{1 - \alpha \beta_N}. \quad (9)$$

The parameter μ captures the local expenditure channel. It increases when workers spend more of their income locally and when labor receives a larger share of nontradable revenue.

Local labor supply determines how these demand shifts divide between employment and wages. Workers choose locations based on wages and local living costs. We write labor supply to sector $j \in \{S, N\}$ as

$$L_j^s = \bar{L}_j \left(\frac{w_j}{R^\chi} \right)^{\varepsilon_j}, \quad \varepsilon_j > 0, \quad (10)$$

where R denotes local housing costs, ε_j governs the response of local labor supply, and χ captures the importance of housing costs for workers' location decisions. This formulation also allows workers to commute into the local economy. When commuting is common, local housing costs play a smaller role in labor supply.

An expanding local workforce also raises housing demand. Housing supply satisfies

$$H^s = \bar{H} R^{\eta_H}, \quad \eta_H > 0, \quad (11)$$

where η_H governs how readily housing supply responds to higher local housing costs. When housing supply responds slowly, more of the increase in local demand appears in R rather than new housing.

These conditions describe how the local economy absorbs an increase in computing capacity. A larger K changes direct employment through equation (6) and shifts linked-sector labor demand through equation (7). The resulting income increases demand for nontradables through equation (9). Wages, employment, and local housing costs adjust in response.

Around the initial equilibrium, we summarize these adjustments as

$$\hat{L}_j = \Theta_j \hat{K}, \quad j \in \{S, N\}, \quad (12)$$

where Θ_j combines the underlying demand response with the adjustment of local labor and housing markets. This expression corresponds to the reduced-form relationship in equation (2). The coefficients Θ_S and Θ_N need not be the same across sectors. Production linkages can generate a large response in S even when the local expenditure response in N is small.

A.2.3 Comparative Statics and Interpretation

The model shows how a labor-light investment can generate broader local employment. A larger K can increase direct data center employment, but employment need not rise proportionally with capacity because employment intensity declines with scale. Data centers also increase demand for local inputs and services. The more of these inputs they purchase locally, the larger the employment response in the linked sector. Large employment gains can therefore occur outside the facilities themselves.

The employment response can spread beyond firms that supply data centers. New jobs in D and S raise local labor income. Workers spend some of that income at restaurants, retailers, and other local businesses, increasing demand for labor in N . Equation (9) captures this feedback. The more income workers spend locally, the larger the nontradable response.

Local markets determine how these demand shifts translate into employment, wages, and local costs. Firms can hire additional workers or bid up wages. An expanding workforce also increases housing demand. Labor mobility and local housing supply determine how the economy accommodates this growth. The equilibrium coefficients Θ_S and Θ_N summarize these adjustments and the resulting employment response.

The sectoral composition of employment helps distinguish the underlying channels. Employment growth in industries that supply complementary inputs and services points to local production linkages. Employment growth in restaurants, retail, and other nontradable services points to a local expenditure channel. Together, these responses show how data center investment can spread through the broader local economy beyond a capital-intensive facility.

However, employment growth is not a measure of local welfare. Data center investment can raise employment and labor income while also increasing housing, infrastructure, and other local costs.

Who ultimately gains depends on who works locally, who owns property, and who bears these costs. We therefore distinguish the effects of data centers on local economic activity from their effects on local welfare.

A.2.4 Dynamic Adjustment

Data center investment unfolds over time. A project moves from construction to activation and, in many cases, subsequent expansion. These stages generate different local effects. Construction creates an immediate but temporary demand for labor. Activation begins ongoing operating employment in D and starts the production and expenditure channels described above. Employment in S and N may adjust more gradually as local firms expand and workers enter the local economy.

Installed capacity can also change after the initial facility opens. Let ΔK_{t+1}^I denote capacity added by incumbent operators and ΔK_{t+1}^E capacity added by new entrants. Abstracting from retirements, installed capacity evolves according to

$$K_{t+1} = K_t + \Delta K_{t+1}^I + \Delta K_{t+1}^E, \quad (13)$$

where K_t denotes installed capacity at time t . Incumbent operators may expand existing campuses in phases. Subsequent capacity from an incumbent can therefore reflect continued investment at an existing site rather than attraction of a new operator. Entry by a new operator provides stronger evidence that existing data center activity makes the location attractive for additional investment.

Existing capacity may make a location more attractive to new operators. Data centers can benefit from shared power and network infrastructure, specialized local suppliers, or proximity to other computing facilities. We write expected investment by new entrants as

$$\mathbb{E}_t [\Delta K_{t+1}^E] = g(K_t, Z_t), \quad (14)$$

where Z_t collects other local characteristics that affect data center investment. Colocation forces imply

$$\frac{\partial g(K_t, Z_t)}{\partial K_t} > 0. \quad (15)$$

Existing capacity then predicts subsequent investment by new operators.

The distinction between initial entry and subsequent capacity also matters for local employment. The first facility starts the local adjustment process, while later additions to K create additional

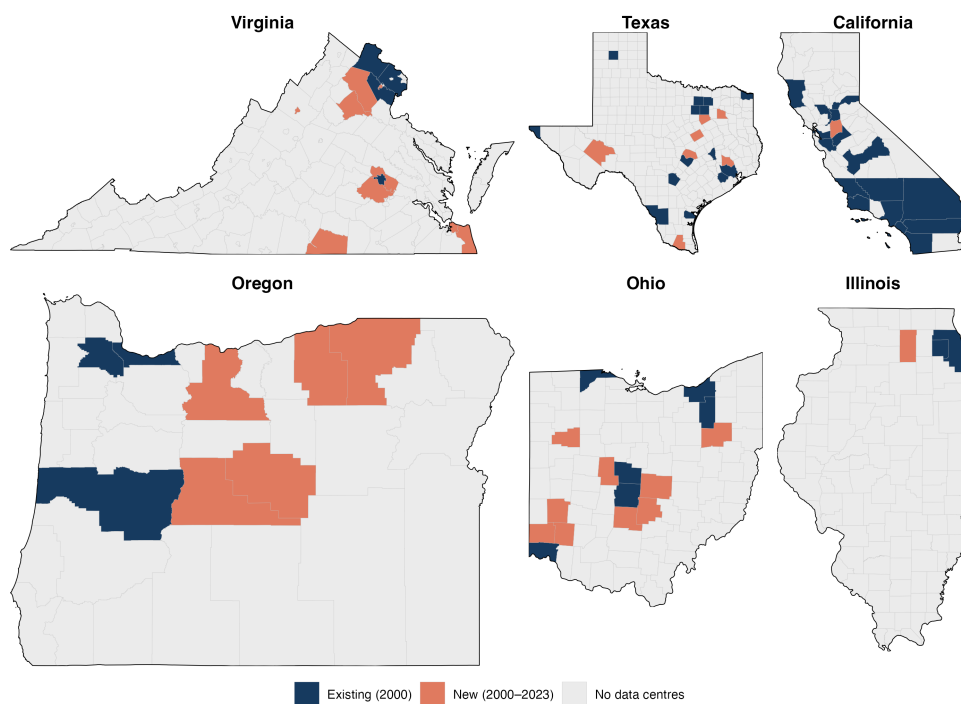
demand for workers, suppliers, and local services. An event study centered on the first facility therefore combines the response to initial entry with any response to subsequent capacity expansion.

A.3 Data

A.3.1 Geographic Expansion

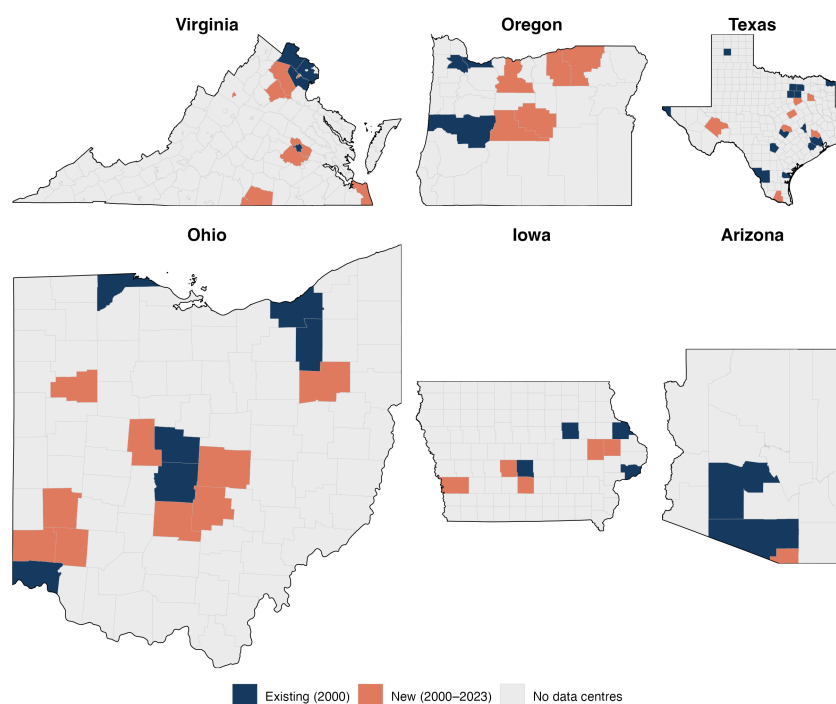
Figures A1 and A2 provide a closer view of the geographic expansion documented in the main text. Figure A1 focuses on the six states with the largest increase in the number of active data centers between 2000 and 2023. Figure A2 instead selects the six states with the largest increase in installed capacity. Both figures show that data center growth expanded beyond counties that already hosted facilities in 2000.

Figure A1: **Geographic Expansion of Data Center Activity, 2000–2023**



Notes: The figure shows the geographic expansion of data center activity between 2000 and 2023 across the six states with the largest increases in the number of active data centers during this period. Dark-shaded counties hosted at least one active data center in 2000 that remained active in 2023. Orange-shaded counties had no active data center in 2000 but hosted at least one by 2023. Light-shaded counties had no active data center in either year.

Figure A2: **Geographic Expansion of Data Center Activity in States with the Largest Capacity Growth, 2000–2023**

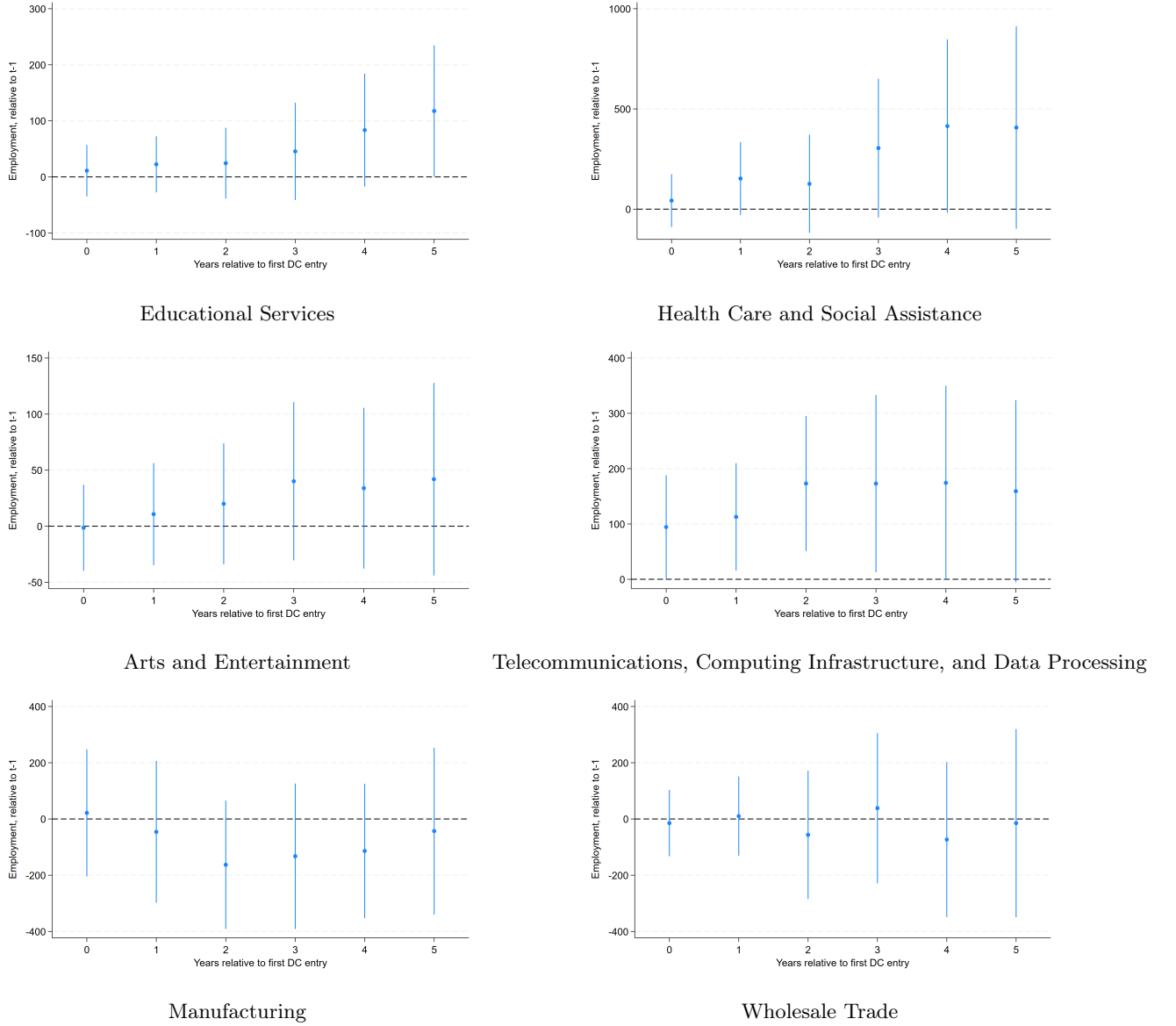


Notes: The figure shows the geographic expansion of data center activity between 2000 and 2023 across the six states with the largest increases in installed data center capacity (measured in MW) over this period. Dark-shaded counties hosted at least one active data center in 2000 that remained active in 2023. Orange-shaded counties had no active data center in 2000 but hosted at least one by 2023. Light-shaded counties had no active data center in either year.

A.4 Additional Extensive-Margin Results

A.4.1 Additional Sector Results

Figure A3: Employment Effects of First Data Center Activation in Additional Sectors



Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates for sector employment following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The donor pool consists of counties that never receive a data center during the sample period. The same SDiD unit and time weights estimated for non-construction employment are used for all sector outcomes. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Vertical bars report 95% jackknife confidence intervals.

Figure A3 extends the sectoral analysis to additional industries. Employment in educational services and health care and social assistance increases at longer horizons. We also examine the industries within the information sector most closely related to data center activity: telecommunications, computing infrastructure, and data processing (NAICS 517 and 518). Employment in these industries increases after activation and stabilizes at just below 200 additional jobs by the second year.

A.4.2 Employment Results Using QWI: Extensive Margin

We re-estimate the extensive-margin employment results using the Census Quarterly Workforce Indicators (QWI). The QWI provide an alternative measure of county-level employment based primarily on state unemployment insurance wage records. Coverage is more limited than in CBP because some workers are not covered by unemployment insurance and the Census suppresses county-industry cells when disclosure is a concern. We do not impute these suppressed observations.

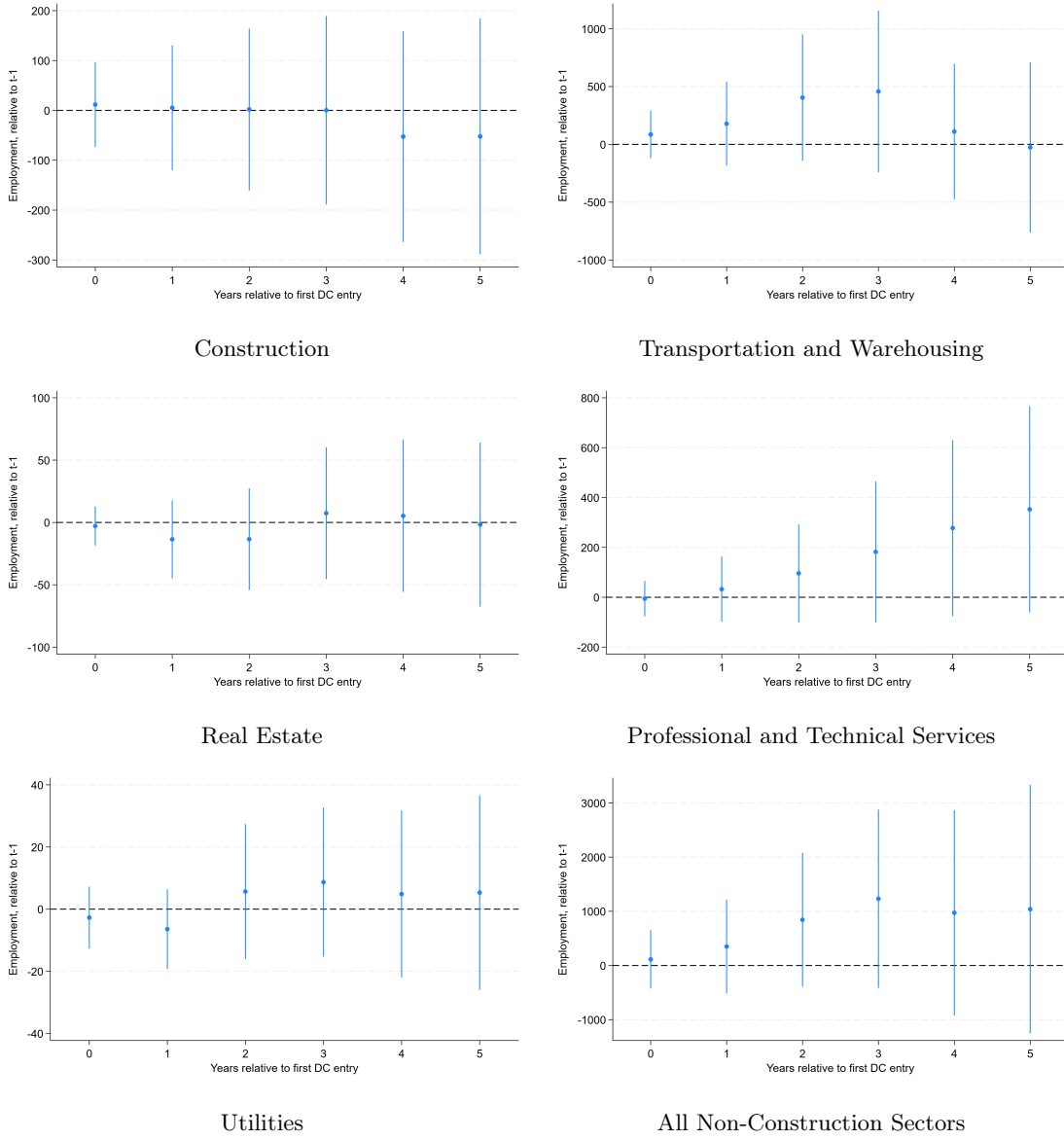
Figure A4 reports the results for non-construction employment and the sectors most closely linked to data center activity. The QWI estimates broadly confirm the patterns in the CBP data. Total non-construction employment increases gradually after the first data center activation. Professional and technical services show a particularly strong increase at longer horizons. Real estate also shows positive point estimates, although they are imprecise. Transportation and warehousing shows little clear response, as do utilities.

Figure A5 examines other local sectors. Retail trade shows little evidence of employment growth. In contrast, accommodation and food services, information, and administrative services show positive point estimates at longer horizons, although the confidence intervals are generally wide.

Figure A6 extends the analysis to additional sectors. Educational services, health care and social assistance, and arts and entertainment show positive employment responses at longer horizons. Manufacturing and wholesale trade show no clear response.

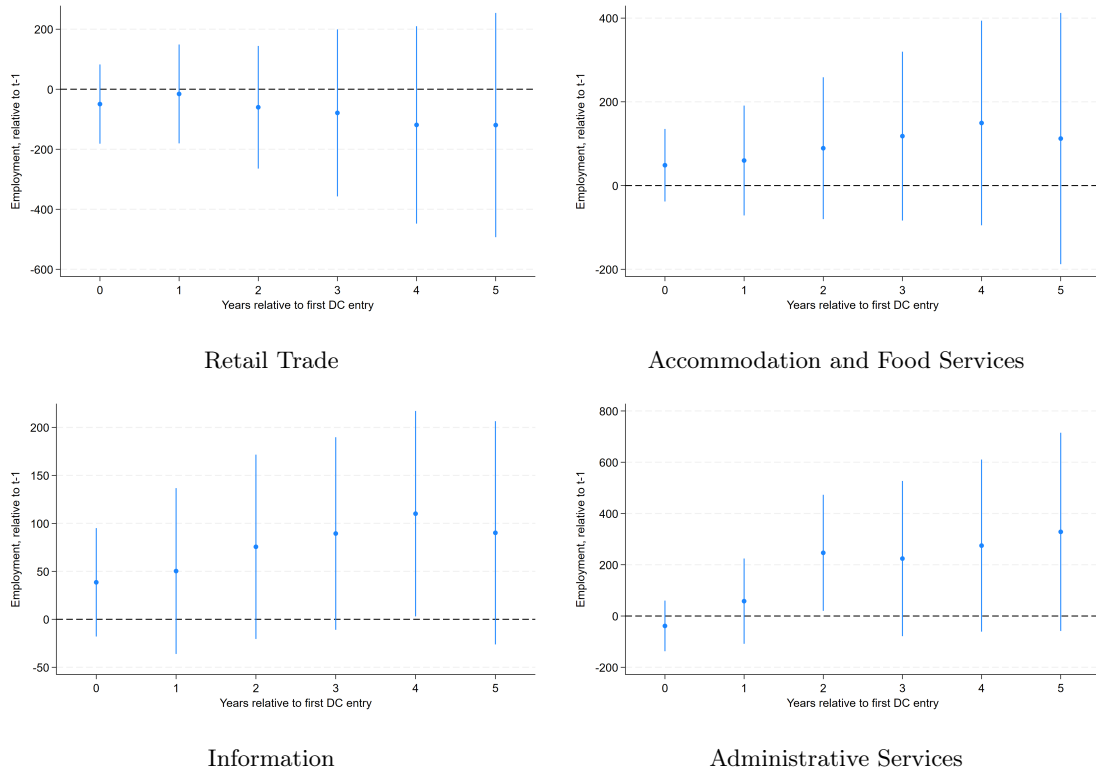
Overall, the QWI estimates support the main finding that non-construction employment increases following the first data center activation. They also show positive point estimates in several sectors linked to data center activity, particularly professional and technical services and real estate. The sectoral estimates are less precise than in CBP and provide some evidence of employment growth in additional local service sectors. We therefore view the QWI results as broadly supportive of the aggregate employment response, while the detailed sectoral composition is less sharply estimated.

Figure A4: **Employment Effects of First Data Center Activation by Selected Sector, QWI**



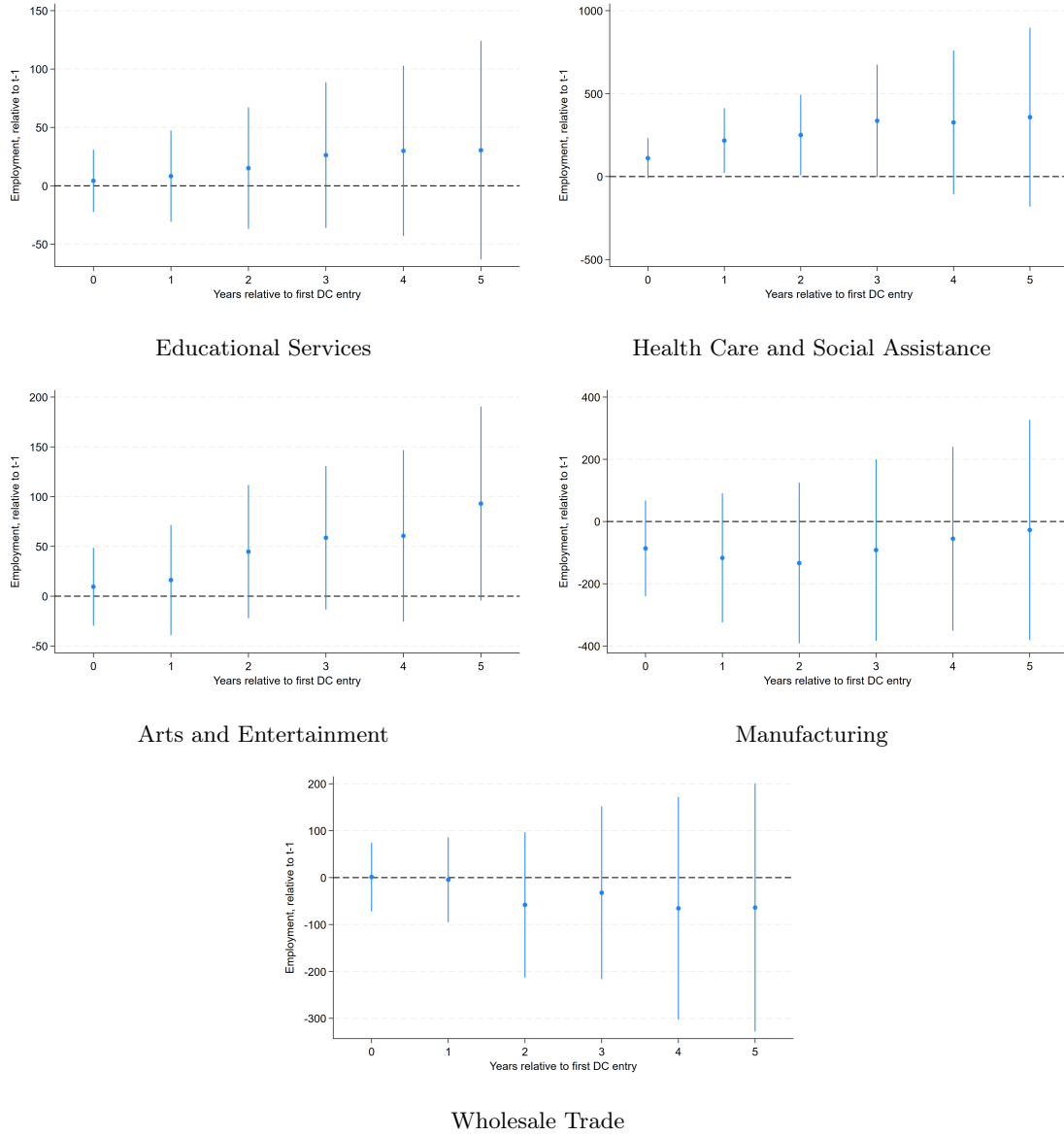
Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates using employment from the Quarterly Workforce Indicators (QWI) following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. For construction, employment is shifted back three years relative to the first data center activation to capture activity associated with building the facility before it becomes operational. For all other sectors, event time is defined relative to the year of activation. The donor pool consists of counties that never receive a data center during the sample period. The same SDiD unit and time weights estimated for QWI non-construction employment are used for all sector outcomes. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Vertical bars report 95% jackknife confidence intervals.

Figure A5: **Employment Effects of First Data Center Activation in Other Local Sectors, QWI**



Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates using employment from the Quarterly Workforce Indicators (QWI) following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The donor pool consists of counties that never receive a data center during the sample period. The same SDiD unit and time weights estimated for QWI non-construction employment are used for all sector outcomes. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Vertical bars report 95% jackknife confidence intervals.

Figure A6: **Employment Effects of First Data Center Activation in Additional Sectors, QWI**



Notes: The figure reports dynamic synthetic difference-in-differences (SDiD) estimates using employment from the Quarterly Workforce Indicators (QWI) following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The donor pool consists of counties that never receive a data center during the sample period. The same SDiD unit and time weights estimated for QWI non-construction employment are used for all sector outcomes. All estimates use outcomes residualized by county and state-by-year fixed effects and the controls described in Section 4.1.1. Vertical bars report 95% jackknife confidence intervals.

A.4.3 Local Projection Difference-in-Differences

Specification. We also estimate the extensive-margin results using the LP-DiD estimator of [Dube et al. \(2025\)](#). This provides an alternative estimate of the response to a county’s first data center activation. LP-DiD compares newly treated counties with counties that have not yet been treated. Unlike SDiD, it relies on conditional parallel trends rather than reweighting control counties to match the pre-treatment outcome paths of treated counties. We therefore use LP-DiD as a robustness check rather than our preferred extensive-margin estimator.

Let D_{ct} equal one if county c has an operational data center by year t , and let $\Delta D_{ct} = D_{ct} - D_{c,t-1}$. Thus, ΔD_{ct} equals one in the year of the county’s first data center activation. Following [Dube et al. \(2025\)](#), we estimate

$$y_{c,t+h} - y_{c,t-1} = \beta^h \Delta D_{ct} + \delta_{st}^h + \omega_c^h + \mathbf{x}_{c,t}' \boldsymbol{\gamma}^h + \varepsilon_{ct}^h, \quad (16)$$

for horizons $h = -5, \dots, 5$. The coefficient β^h compares employment at $t + h$ with employment in the year before activation. For $h \geq 0$, it traces the employment response following activation. For $h \leq -2$, it compares earlier pre-treatment outcomes with the same $t - 1$ reference year. We normalize $h = -1$ to zero.

For each post-treatment horizon, we keep as controls only counties that remain untreated through $t+h$. This avoids using counties that activate a data center during the estimation window as controls. As in the baseline analysis, we exclude the top 5% of counties by employment in 2000. We also exclude counties that already hosted a data center before 2000 because we do not observe their pre-treatment period.

The regressions include county and state-by-year fixed effects. We also include Republican vote share and baseline county characteristics interacted with linear year trends. These characteristics are measured in or prior to 2000 and include industrial electricity prices, internet infrastructure, employment growth from 1980 to 2000, educational attainment, the shares of foreign-born and young residents, water coverage, January temperature, and scenic amenity. Standard errors are clustered by county.

Results. Figure [A7](#) shows the LP-DiD estimates for non-construction employment. The aggregate pre-activation estimates suggest a small pre-trend with employment increasing to approximately 4,000 additional jobs five years after activation. This post-activation response is substantially larger

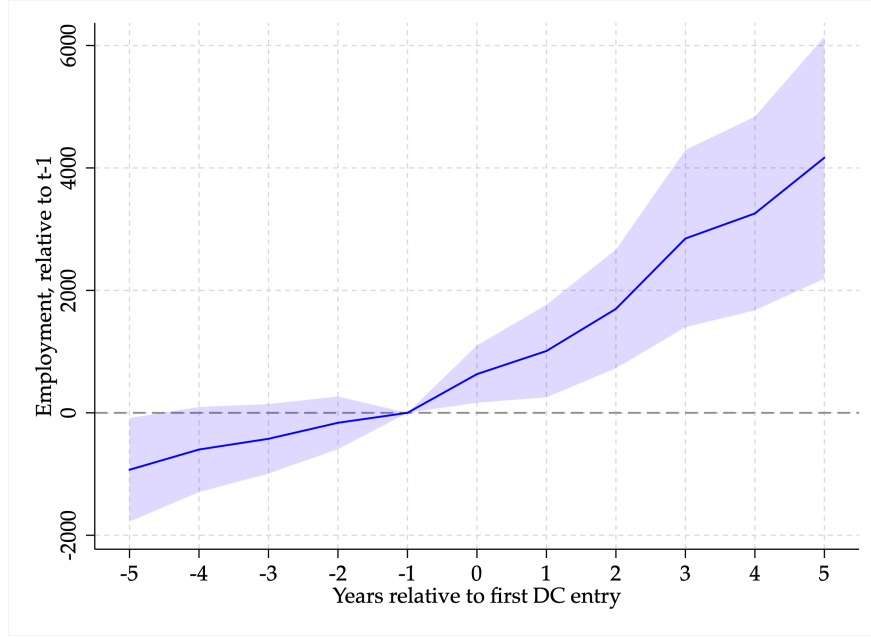
than the corresponding SDiD estimates but does suffer from pre-trends.

Figure A8 reports the sectoral results for industries more closely linked to data center activity. Real estate and professional and technical services increase after activation. Transportation and warehousing also shows a positive post-activation response, but its pre-treatment estimates indicate that treated counties were already on a different trajectory before the first data center opened. Construction and utilities remain imprecisely estimated.

Figure A9 shows stronger responses in local service sectors than the SDiD estimates. Retail trade, accommodation and food services, and administrative services increase after activation. Information also shows positive point estimates at longer horizons. For retail trade and accommodation and food services, however, the pre-activation estimates also provide evidence of differential trends before treatment.

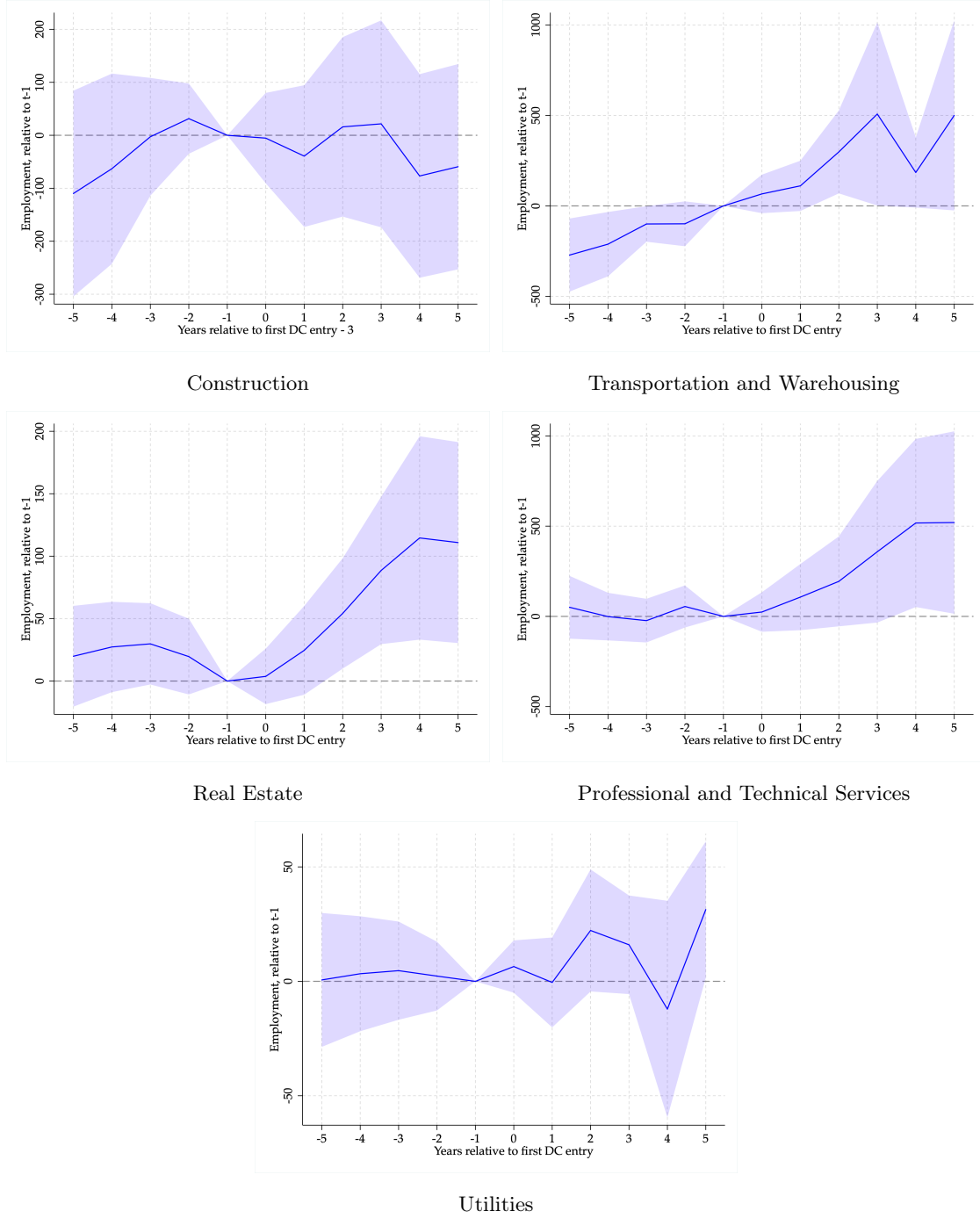
The LP-DiD results therefore require a more cautious interpretation. Aggregate employment increases gradually after activation, but the estimated response is considerably larger than under SDiD. Several sectoral outcomes also show evidence that treated counties were on steeper trajectories before data center entry. This pattern can contribute to the broader and larger post-treatment responses estimated by LP-DiD. SDiD addresses this concern by reweighting control counties to more closely match the pre-treatment outcome paths of treated counties. We therefore view SDiD as our preferred extensive-margin estimator and use LP-DiD as a robustness check on the aggregate employment response.

Figure A7: **Non-Construction Employment Following First Data Center Activation, LP-DiD**



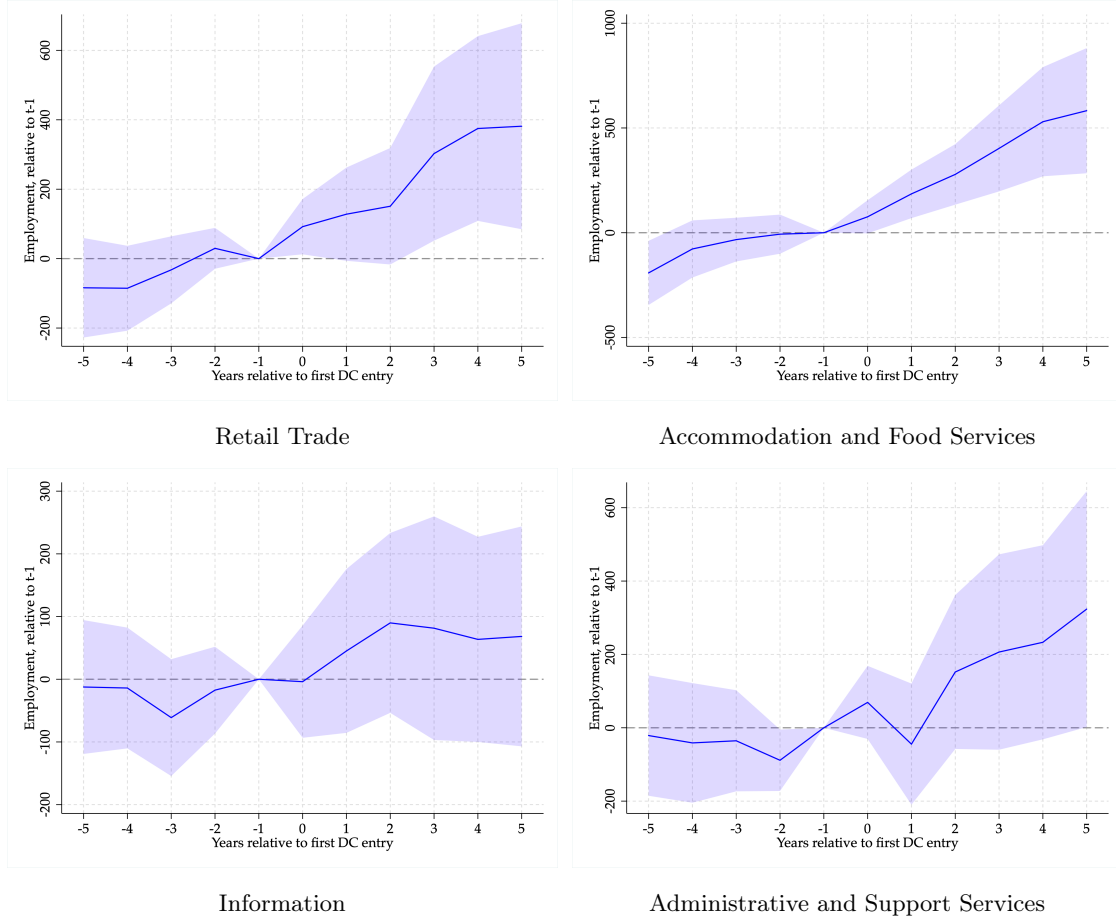
Notes: The figure reports local projection difference-in-differences (LP-DiD) estimates of the employment response following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The figure reports estimates for event-time horizons $h = -5, \dots, 5$. For post-treatment horizons, the control group consists of counties that remain untreated through the relevant horizon. All specifications include county and state-by-year fixed effects and the controls described in Section A.4.3. Shaded bands report 95% confidence intervals based on standard errors clustered by county.

Figure A8: **Employment Effects of First Data Center Activation by Selected Sector, LP-DiD**



Notes: The figure reports local projection difference-in-differences (LP-DiD) estimates for selected sectors following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The figure reports estimates for event-time horizons $h = -5, \dots, 5$. For post-treatment horizons, the control group consists of counties that remain untreated through the relevant horizon. All specifications include county and state-by-year fixed effects and the controls described in Section A.4.3. Shaded bands report 95% confidence intervals based on standard errors clustered by county.

Figure A9: **Employment Effects of First Data Center Activation in Other Local Sectors, LP-DiD**



Notes: The figure reports local projection difference-in-differences (LP-DiD) estimates for selected sectors following a county's first data center activation. Year 0 denotes the first year in which the data center is operational, and estimates are relative to year $t - 1$. The figure reports estimates for event-time horizons $h = -5, \dots, 5$. For post-treatment horizons, the control group consists of counties that remain untreated through the relevant horizon. All specifications include county and state-by-year fixed effects and the controls described in Section A.4.3. Shaded bands report 95% confidence intervals based on standard errors clustered by county.

A.5 Additional Intensive-Margin Results

A.5.1 Employment Spillovers to Adjacent Counties

The employment gains in host counties could reflect relocation from nearby counties rather than new regional employment. We test this possibility using the Census county adjacency file. For each focal county, we sum non-construction employment across all directly adjacent counties and re-estimate our preferred specification using this neighboring-county employment measure as the

outcome.

Table A1 reports the results. We find no evidence that employment falls in neighboring counties as data center capacity expands. The estimated coefficient is positive in both specifications, at 5.34 jobs per MW in the baseline sample and 17.66 jobs per MW when we additionally exclude Loudoun County. Neither estimate is statistically significant. The results provide no evidence that employment gains in host counties come at the expense of directly adjacent counties, but they do not establish positive spillovers to neighboring counties.

Table A1: **Employment Spillovers to Adjacent Counties**

	(1) Baseline sample	(2) Excl. Loudoun
Installed DC capacity (MW), year t	5.344 (6.610)	17.66 (12.90)
Observations	52,128	52,110

Notes: The table reports estimates of equation (5) relating installed data center capacity in the focal county to employment in adjacent counties. The dependent variable is the cumulative change in the sum of non-construction employment across all counties directly adjacent to the focal county from $t - 1$ to $t + 5$. Adjacent counties are identified using the Census county adjacency file. The baseline sample excludes the top 5% of counties by employment in 2000. Column (2) additionally excludes Loudoun County, Virginia. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1. Standard errors clustered by focal county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.5.2 Sensitivity to Large Data Center Hubs

Table A2 examines whether the intensive-margin results depend on the largest data center hubs. We sequentially exclude counties with the highest peak installed capacity from the baseline sample.

The employment coefficient remains positive as we remove increasingly large hubs. It rises from 15.0 jobs per MW after excluding the five largest hubs to 31.5 after excluding the largest 20, before falling to 26.2 after excluding the largest 25. The estimates become less precise as more hubs are removed. The positive point estimates therefore persist outside the largest data center markets.

Follow-on capacity shows a different pattern. The coefficient falls from 0.35 after excluding the five largest hubs to approximately zero after excluding the largest 10. It turns negative once we exclude 15 or more hubs. The positive follow-on capacity response is therefore concentrated in the largest data center markets.

Table A2: Sensitivity to Excluding Large Data Center Hubs

<i>Panel A: Cumulative change in non-construction employment ($t - 1$ to $t + 5$)</i>					
	(1)	(2)	(3)	(4)	(5)
	Excl. top 5	Excl. top 10	Excl. top 15	Excl. top 20	Excl. top 25
Installed DC capacity (MW), year t	14.97** (5.962)	15.15* (8.436)	20.53 (13.42)	31.52** (14.35)	26.16 (21.06)
Within R^2	0.180	0.179	0.180	0.181	0.183
Observations	52,470	52,380	52,290	52,200	52,110
<i>Panel B: Cumulative change in installed data center capacity (t to $t + 6$)</i>					
	(1)	(2)	(3)	(4)	(5)
	Excl. top 5	Excl. top 10	Excl. top 15	Excl. top 20	Excl. top 25
Installed DC capacity (MW), year t	0.349* (0.184)	0.006 (0.174)	-0.223* (0.122)	-0.119** (0.0521)	-0.121** (0.0584)
Within R^2	0.062	0.019	0.042	0.030	0.032
Observations	52,470	52,380	52,290	52,200	52,110

Notes: The table examines the sensitivity of the intensive-margin estimates to excluding the largest data center hubs from the baseline sample. The baseline sample excludes the top 5% of counties by employment in 2000. Among the remaining counties, data center hubs are ranked by their maximum installed capacity between 2001 and 2023. Columns (1)–(5) sequentially exclude the 5, 10, 15, 20, and 25 counties with the highest maximum installed capacity. Panel A reports the cumulative change in non-construction employment from $t - 1$ to $t + 5$ associated with one additional MW of installed data center capacity in year t . Panel B reports subsequent installed data center capacity from t to $t + 6$ associated with one additional MW of installed capacity in year t . All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1. Standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

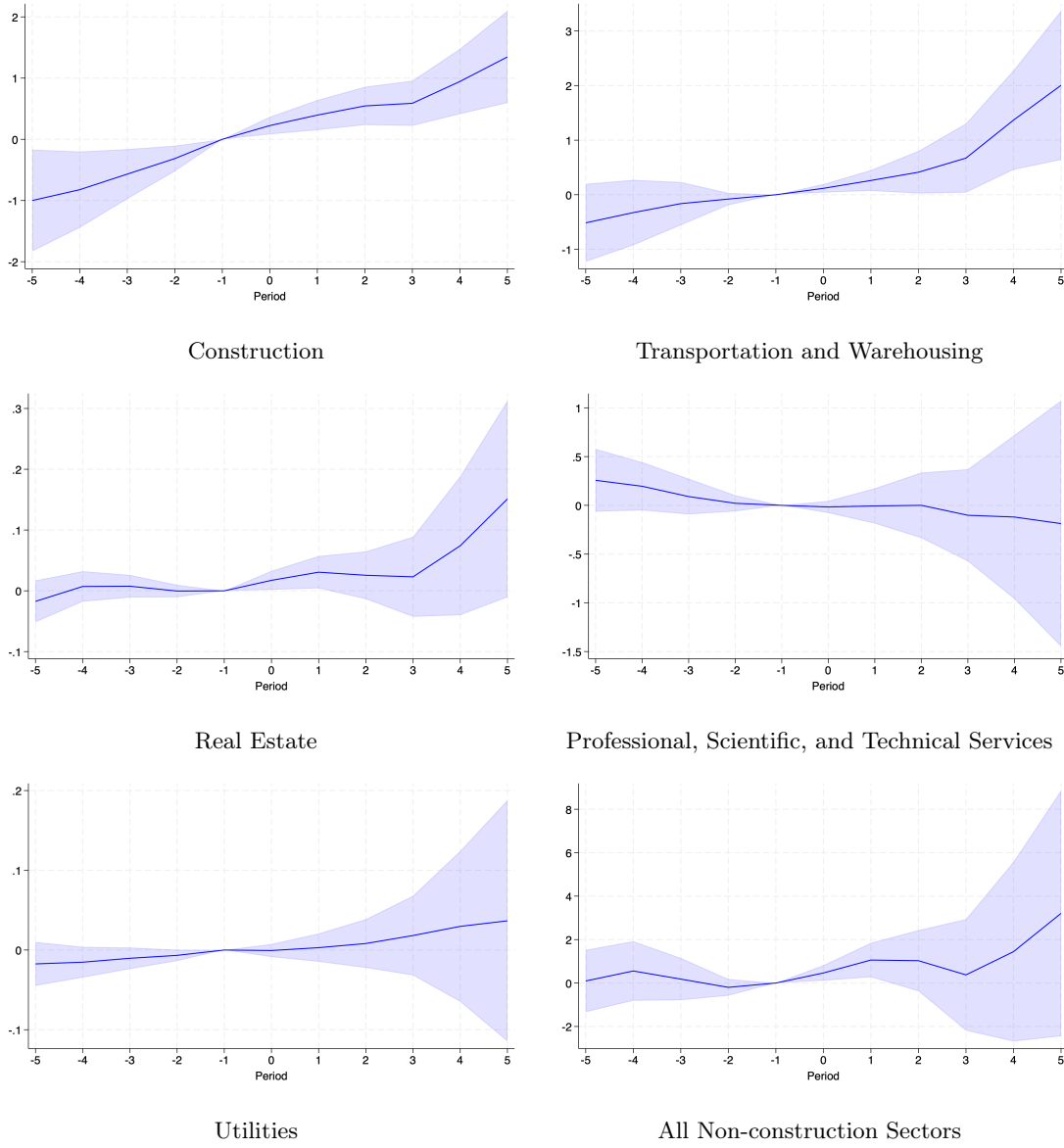
A.5.3 Employment Results Using QWI: Intensive Margin

Figure A10 shows the results for the sectors with the strongest employment responses in the CBP data. The QWI estimates broadly confirm the main pattern. Employment rises in transportation and warehousing, professional and technical services, and total non-construction employment. The estimates become less precise at longer horizons, consistent with the QWI data's more limited coverage.

Figure A11 reports the remaining sectors. As in the CBP data, we find little evidence of employment growth in retail trade. Information and administrative services also do not increase. Accommodation and food services show a positive response at longer horizons, although the estimates become imprecise.

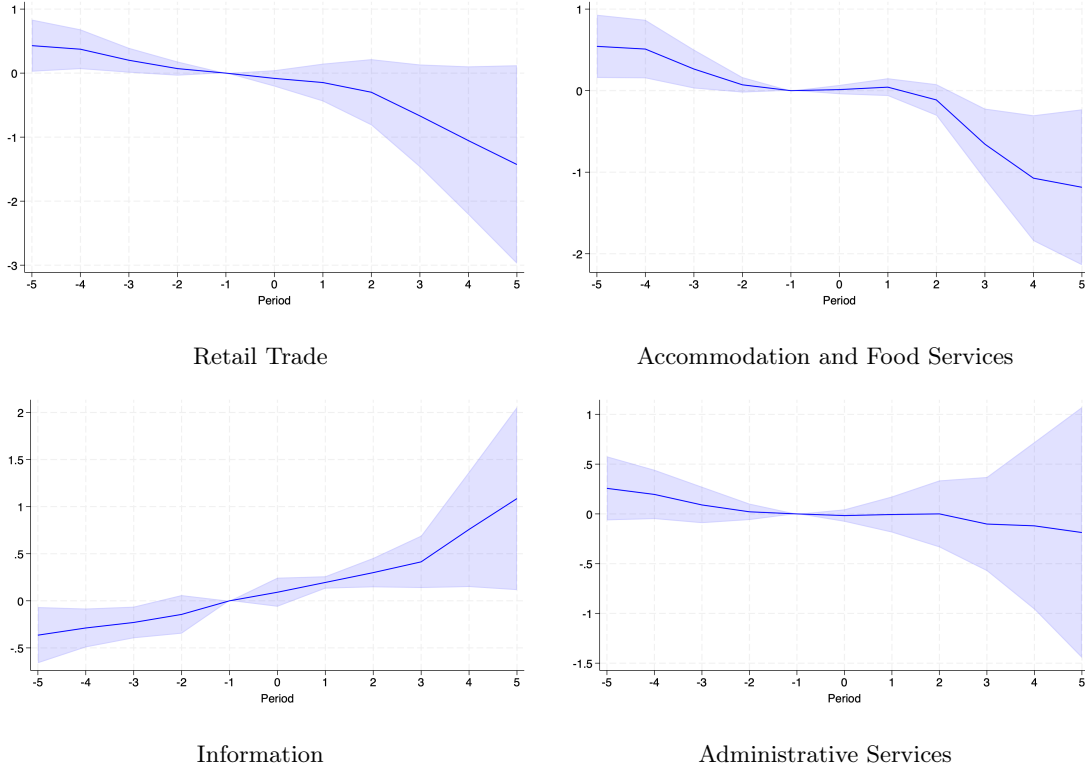
Overall, the QWI results reinforce the main conclusion from the CBP data. Employment gains are concentrated in a narrow set of sectors linked to data center activity rather than spread broadly across the local economy.

Figure A10: **Employment Effects per MW by Selected Sector, QWI**



Notes: The figure reports estimates of equation (5) using employment from the Quarterly Workforce Indicators (QWI). Each panel shows the cumulative change in employment from $t - 1$ through horizon h associated with one additional MW of installed data center capacity in county c and year t . The final panel reports total non-construction employment in the QWI. Shaded bands report 95% confidence intervals based on standard errors clustered by county. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1.

Figure A11: **Employment Effects per MW in Other Local Sectors, QWI**



Notes: The figure reports estimates of equation (5) using employment from the Quarterly Workforce Indicators (QWI). Each panel shows the cumulative change in employment from $t - 1$ through horizon h associated with one additional MW of installed data center capacity in county c and year t . Shaded bands report 95% confidence intervals based on standard errors clustered by county. All specifications include county and state-by-year fixed effects and the controls described in Section 4.2.1.