OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Data Center Development and Local Economy: Assessing the Persistence of Economic Growth Benefits

Travyn Rasiah
travyn.rasiah@gmail.com

ABSTRACT

This paper examines whether data centers generate sustained local economic growth or whether their effects are short-lived and limited. As demand for cloud computing, artificial intelligence, and digital infrastructure continues to rise, data centers have become increasingly important to regional economic development. Using county-level GDP data from the U.S. Bureau of Economic Analysis from 2001 to 2024, this study analyzes the relationship between major data center activity and local economic output. The analysis uses a two-way fixed effects regression with log GDP as the dependent variable and data center treatment as the main independent variable, while controlling for county-specific and year-specific factors. The results show that data center treatment has a positive and statistically significant relationship with county-level GDP. The coefficient of 0.1609 translates to an estimated 17.46% increase in GDP for treated counties. However, the findings do not prove that data centers alone caused all of this growth. The regression also pools its estimate across all post-treatment years, so it cannot tell whether gains build over time, hold steady, or diminish after the initial investment. Instead, data centers appear to function as one contributor to broader growth-related conditions, such as infrastructure investment, business activity, and regional technology development. The evidence points toward data centers as contributors to local economic growth, though their long-term impact depends on surrounding economic conditions and whether communities can channel that initial investment into broader development.

KEYWORDS
Data centers, Local economic growth, County-level GDP, Gross Domestic Product, Log GDP, County fixed effects, Year fixed effects, Tech Treatment, Treated counties, Control counties

INTRODUCTION

Data centers have become one of the most important pieces of the modern economy, even though most people rarely think about them directly. Every time people use cloud storage, stream video, search online, run a business platform, or use artificial intelligence tools, some part of that activity depends on data centers. Because of this, counties and cities across the United States have started competing to attract major facilities from companies such as Amazon Web Services, Microsoft, and Google. Local governments often view these projects as signs of future growth, especially because they bring large investments in land, construction, power systems, and digital infrastructure.

At the same time, the actual local impact of data centers is not as simple as it may appear. Supporters argue that data centers can raise local GDP, expand the tax base, improve infrastructure, and attract related technology activity. Critics point out that many data centers create relatively few permanent jobs after construction ends, use large amounts of electricity and water, and often receive tax incentives that reduce the public benefits communities expect to gain. This makes the question more complicated than simply asking whether data centers are good or bad for local economies.

This paper asks whether data centers generate sustained local economic growth or whether their effects are short-lived and limited. To study this, the paper uses county-level GDP data from the U.S. Bureau of Economic Analysis from 2001 to 2024 and compares counties with major data center activity to counties without major data center treatment. The analysis uses a two-way fixed effects regression with log GDP as the dependent variable, allowing the model to control for stable differences between counties and broader economic changes across years.

The main purpose of this study is not to claim that data centers alone explain local economic growth. Instead, it asks whether data center presence is associated with stronger county-level GDP after accounting for county and year effects. The results show a positive and statistically significant relationship between data center treatment and local GDP. This finding is consistent with sustained growth, but the model averages across all post-treatment years and cannot directly test whether that effect holds, grows, or fades.

The rest of the paper proceeds as follows. Section 2 reviews the existing literature on data centers and local economic development. Section 3 describes the dataset and summary statistics. Section 4 explains the two-way fixed effects methodology. Section 5 presents the regression findings. Section 6 discusses the interpretation and limitations of the results. Section 7 concludes.

LITERATURE REVIEW

The rapid expansion of hyperscale data centers has made digital infrastructure one of the more important and still understudied drivers of regional economic change. As artificial intelligence and cloud computing continue to grow, researchers and policymakers have become more interested in how data centers affect local economic output. However, the existing literature is still limited and uneven.

Much of the most cited evidence on the macroeconomic impact of data centers comes from simulation-based studies. PwC (2024) uses IMPLAN input-output modeling to estimate that U.S. data centers generated substantial gains in GDP, labor income, and employment between 2017 and 2021, with the sector growing faster than the broader economy. These estimates provide a useful benchmark, but they should be interpreted with caution. Input-output models track how spending moves through the economy rather than measuring actual changes in regional output, and they can overstate multiplier effects that may not fully materialize in practice. The PwC (2024) report is helpful for understanding the scale of data center-related activity, but it does not show whether data center entry causes lasting changes in local economic conditions.

Studies from outside the United States add to this picture, though cross-country comparability remains uncertain. Li, He, and Cao (2025), in a panel-data study of green data center pilots in China, find statistically significant effects on urban digital economic development, with benefits spilling over into nearby regions. This finding supports the idea that data centers may serve as anchors for broader technology clusters, rather than operating simply as isolated facilities. However, the applicability of these results to American counties is uncertain because China's industrial policy, state-directed investment, and urban institutions differ from those in the United States.

At the local level, the picture is more complicated. Muro, Parilla, and Maxim (2026) offer a more measured view. They find that data centers can bring tax revenue and infrastructure investment, especially to rural communities, but their employment effects are often limited and their broader economic benefits vary widely from place to place. One reason is that large facilities are frequently attracted through negotiated tax incentives. While these incentives may help communities win projects, they can also reduce the fiscal benefits that local governments originally expected (Muro, Parilla, & Maxim, 2026). In some cases, data centers operate as capital-intensive enclaves with few durable ties to the regional economy, which raises doubts about their ability to maintain GDP growth beyond the construction phase.

Evidence from the firm level rounds out the picture. Park (2025) finds that proximity to data centers is associated with measurable improvements in firm productivity, sales, and profitability. The mechanism is plausible: reduced latency, better access to cloud infrastructure, and lower data processing costs could improve business performance (Park, 2025). However, the study stops short of tracing these firm-level gains to aggregate regional output. Whether productivity effects at the firm level scale up to drive county-level GDP remains an open empirical question.

One dimension largely absent from GDP-focused analyses is the cost side. Ngata et al. (2025) document the energy and water demands that large data centers impose on local infrastructure. These pressures impose genuine economic costs even if they are not fully visible in output statistics. Standard GDP measures do not capture these externalities, implying that the current studies may present an incomplete picture of the overall welfare effects of data center expansion.

Across these studies, data centers appear to enhance economic activity through direct investment, indirect supply chain effects, and productivity spillovers (PwC, 2024; Park, 2025). What it does not fully provide is a clear estimate of how data center activity relates to real GDP in U.S. counties over time. Many studies rely on modeling assumptions, non-U.S. contexts, or cross-sectional variation that cannot fully separate the effect of data center presence from the characteristics of the places that attract them. This paper addresses that gap by using a two-way fixed effects regression to compare GDP changes across treated and control counties from 2001 to 2024, while controlling for county-specific and year-specific factors.

DATA AND SUMMARY STATISTICS

This study uses a county-year dataset that I constructed to examine the relationship between data center activity and local economic growth. The main economic outcome is real, inflation-adjusted county-level GDP from the U.S. Bureau of Economic Analysis. The GDP measure uses the "All Industry" total and is reported in thousands of chained dollars. The dataset covers the years 2001 through 2024.

To identify counties affected by data center activity, I classified counties based on whether they contained major facilities operated by companies such as Amazon Web Services, Microsoft Azure, or Google Cloud. Each county-year observation was assigned a binary treatment variable called Tech Treatment. This variable takes on a value of 1 if the county had major data center activity and a value of 0 if the county did not have major data center activity.

The sample consists of 11 U.S. counties and 264 county-year observations. The counties were selected to include both major data center hubs and comparable non-tech counties that serve as a control group. Treated counties include major data center hubs located in states such as Virginia, Washington, Arizona, and Georgia, which are home to hyperscale facilities operated by Amazon Web Services, Microsoft Azure, and Google Cloud. These counties were selected because they represent well-documented cases of major data center investment with publicly available facility opening dates, and they span a range of geographic regions and economic sizes. Control counties were selected to provide geographic and economic diversity without major hyperscale data center presence, and include both rural and suburban counties across different U.S. regions. For each county, total economic growth was measured using annual GDP values from 2001 to 2024, which allows for comparisons across counties with different starting levels of GDP.

Table 1 identifies the counties included in the study, distinguishes treated counties from control counties, and reports the opening year and treatment start year for each qualifying data center facility. Table 2 presents summary statistics for the key variables used in the analysis, including GDP in thousands of chained dollars, log GDP, and the Tech Treatment indicator, reported separately for treated and control county-years.

TABLE 1

County	Status	Treatment Start Year	Post-Treatment Years Through 2024	GDP Before Treatment
Maricopa, AZ	*Treated*	*2011*	*14*	*193,418,704*
Douglas, GA	*Treated*	*2015*	*10*	*3,939,353*
Franklin, OH	*Treated*	*2016*	*9*	*91,913,420*
Dallas, TX	*Treated*	*2010*	*15*	*189,640,088*
Loudoun, VA	*Treated*	*2010*	*15*	*18,586,952*
Fairfax, VA	*Treated*	*2012*	*13*	*104,106,773*
Jefferson, AL	*Control*	—	—	*40,672,335*
Bernalillo, NM	*Control*	—	—	*29,409,858*
Tulsa, OK	*Control*	—	—	*31,696,551*
Milwaukee, WI	*Control*	—	—	*53,673,239*
El Paso, TX	*Control*	—	—	*21,481,244*

TABLE 2

Group	Counties	County-Year Observations	Mean GDP	Standard Deviation of GDP	Mean Log GDP	Standard Deviation of Log GDP
Full sample	***11***	***264***	***80,214,598***	***77,266,882***	***17.711***	***1.097***
Treated counties	***6***	***144***	***112,740,565***	***92,467,212***	***17.885***	***1.447***
Control counties	***5***	***120***	***41,183,438***	***10,176,410***	***17.502***	***0.254***

Because facilities began operating in different years, treated counties contribute different numbers of post-treatment observations through 2024. Counties where major data center facilities opened before 2010 have over a decade of post-treatment GDP data available, which provides some basis for assessing whether growth effects persisted over the medium to long term. Counties where facilities opened after 2015, however, contribute only a limited number of post-treatment years through 2024, making it more difficult to assess whether the observed GDP association reflects a temporary investment effect or a more durable structural change. This variation in treatment timing limits how much the study can say about long-run persistence, particularly for counties where data centers only recently opened.

The summary statistics help show the basic structure of the dataset. They allow the reader to see how GDP differs between treated and untreated counties and how economic output changes over time. This gives a first look at whether data center counties show different growth patterns from control counties.

However, summary statistics alone cannot answer the main research question. A simple comparison between treated and untreated counties could be misleading because counties with data centers may already be different from counties without data centers. For example, data center counties may already have stronger infrastructure, better access to electricity, more available land, larger economies, or better policies. If these pre-existing differences are not accounted for, the paper could incorrectly attribute those advantages to data center activity.

Summary statistics also do not account for year-specific shocks. A recession, inflationary period, or national technology boom could affect many counties at the same time. If GDP rises in treated counties during a national technology expansion, summary statistics alone cannot show whether that increase came from data center activity or from broader economic conditions.

For this reason, the summary statistics are useful as a starting point, but they are not enough to establish the relationship between data center treatment and local GDP. They describe the data and motivate the empirical strategy, but they cannot separate data center treatment from county-specific and year-specific factors. This is why the next section uses a two-way fixed effects regression.

METHODOLOGY

The analysis employs a two-way fixed effects regression, with log GDP as the dependent variable and Tech Treatment as the main independent variable. The model incorporates fixed effects for both counties and years. County fixed effects control for differences between counties, such as geography, infrastructure, population size, and long-term economic conditions. Year fixed effects control for broader economic shocks affecting all counties in the same year, such as recessions, inflation, or growth overall. This approach enables the regression to assess the relationship between data center treatment and GDP while accounting for both county-specific and year-specific factors and influences.

The model can be written as:

$$\log(GDP_{it}) = \beta\ TechTreatment_{it} + \alpha_i + \gamma_t + \varepsilon_{it}$$

In this equation, $\log(GDP_{it})$ represents the natural log of GDP for county i in year t. $TechTreatment_{it}$ is the treatment variable, α_i represents county fixed effects, γ_t represents year fixed effects, and ε_{it} is the error term.

The coefficient on Tech Treatment is the main coefficient of interest. It measures how log GDP changes when a county has data center treatment, after accounting for county fixed effects and year fixed effects. This coefficient is identified by variation in data center treatment within a county over time and by variation in data center treatment across counties within the same year.

This is important because the model is not simply comparing rich counties to poor counties. County fixed effects absorb stable differences between counties, such as whether a county is consistently wealthier, more populated, or better connected to infrastructure. Year fixed effects absorb shocks that affect all counties in the same year, such as recessions, inflation, or national technology booms. As a result, the coefficient estimates how GDP moves with data center treatment after removing both stable county differences and common year-specific shocks.

Several limitations are present in the data collection process. The relatively small sample size reduces statistical reliability and limits generalizability. Additionally, the binary classification of data center presence does not capture variation in the scale or economic impact of different facilities. There is also potential selection bias, as data centers may be located in counties that were already experiencing economic growth. Finally, although the model controls for county and year fixed effects, it does not directly control for other time-varying local factors, such as population growth, employment trends, industry composition, or local policy changes, which may also influence economic outcomes.

The two-way fixed effects model also pools all post-treatment years into a single average, so it cannot tell whether economic gains from data centers build, plateau, or taper off over time. The model cannot distinguish between effects that grow stronger over time, remain stable, or decay after the initial investment period. Directly assessing the persistence of economic effects would require an event study or dynamic treatment effects specification that allows the coefficient to vary by years since treatment. For a paper focused on whether data center effects persist, this is a meaningful gap that future work should address.

FINDINGS

The two-way fixed effects regression found a positively significant relationship between data center treatment and county-level GDP. The dependent variable was log GDP, and the main independent variable was Tech Treatment which is a binary variable that takes a value of 1 if a county had a tech data center that given year and 0 otherwise. The model controlled for both county fixed effects and year fixed effects, meaning it accounted for stable differences between counties and economic changes nationwide across years.

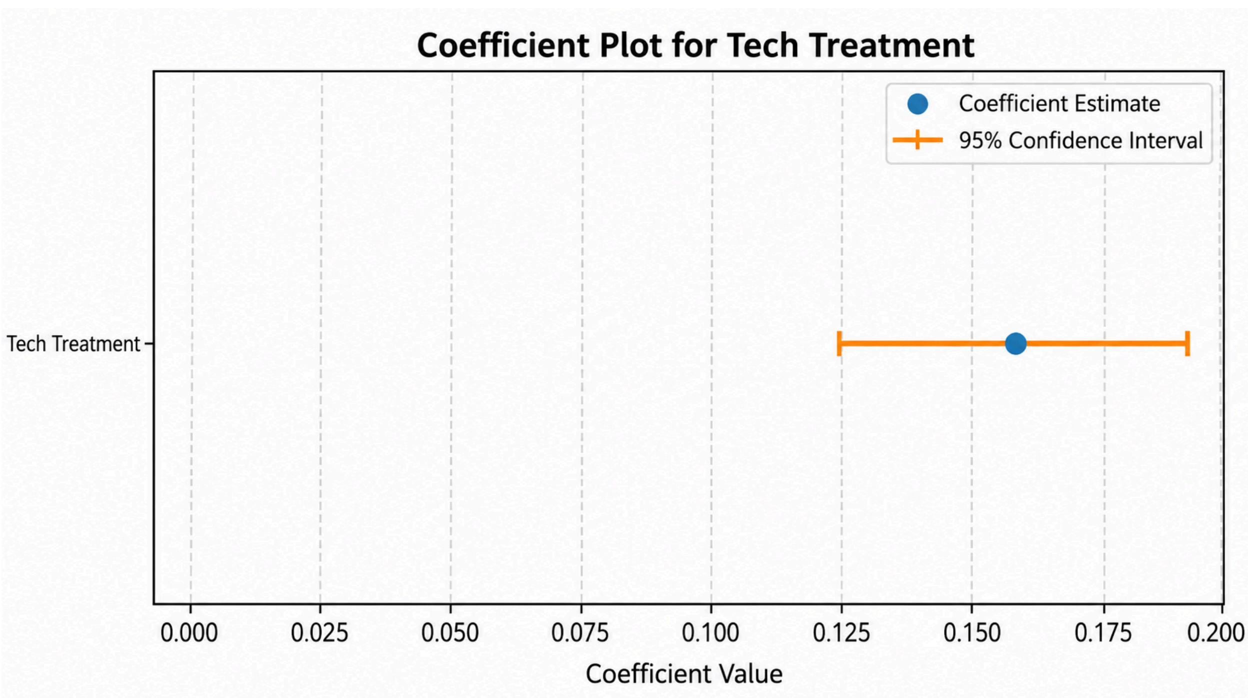

Figure 1: Estimated coefficient for Tech Treatment with a 95% confidence interval. The estimated coefficient is 0.1609, and the confidence interval remains above zero, indicating a positive and significant association between data center treatment and county-level GDP.

The coefficient for Tech Treatment was 0.1609, with a standard error of 0.0181, a t-statistic of 8.9097, and a p-value of 0.0000. Because the p-value is extremely low, the result is statistically significant and suggests that data center treatment is associated with higher local GDP. Since the dependent variable is log GDP, the coefficient can be converted into a percentage effect using $e^{0.1609} - 1$, which equals approximately 17.46%. This means that treated counties were estimated to have about 17.46% higher GDP after data center treatment, holding county-specific and year-specific factors constant.

This suggests that the association between data centers and higher county GDP extends beyond the construction phase, though the model cannot confirm this directly. While the regression does not prove that data centers alone caused the increase in GDP, it does show a strong association between data center treatment and higher county-level economic output after accounting for major county and year-level differences.

The confidence interval also supports this finding. The lower confidence interval was 0.1253, and the upper confidence interval was 0.1965, meaning the estimated effect remains positive across the range of likely values. Overall, the regression results suggest that data center treatment had a significant positive association with county-level GDP.

DISCUSSION

The findings suggest that data centers may contribute to local economic growth, but their effects are not automatically large, permanent, or evenly distributed across counties. Quantitatively, the two-way fixed effects regression shows that data center treatment had a coefficient of 0.1609, a standard error of 0.0181, a t-statistic of 8.9097, and a p-value of 0.0000, meaning the relationship between data center treatment and log GDP is highly statistically significant. Since the dependent variable is log GDP, this coefficient translates to an estimated 17.46% increase in GDP for treated counties, holding county-specific and year-specific factors constant. However, this does not mean that data centers alone caused all of the GDP growth. The regression pools post-treatment years into a single average and cannot separate a lasting structural shift from a temporary burst of investment activity. Instead, because the regression directly measures GDP rather than the specific channels behind GDP growth, the results suggest that data centers may have acted as an economic catalyst by contributing to broader growth-related conditions, such as infrastructure investment, business activity, technology clustering, and regional development.

Counties with major data center activity often show strong long-term GDP growth, especially in areas where data centers are connected to broader technology ecosystems, cloud infrastructure, and business investment. However, the results should not be interpreted as proof that data centers alone caused all of this growth. Many counties that attract data centers already have favorable economic conditions, such as strong infrastructure, business-friendly policies, access to electricity, available land, and proximity to growing metropolitan areas. These existing advantages may make counties more likely to attract data centers in the first place, while data centers may then reinforce and accelerate the growth already happening in those regions.

This matters because data centers are very capital intensive. They require large investments in construction, land, energy systems, and digital infrastructure, which can boost local GDP temporarily. However, once construction is complete, the long-term employment impact may be limited because operating the center does not require as much labor as many other types of centers. This supports the idea that some of the direct economic benefits may be strongest in the short run, especially during the construction and development phase.

At the same time, data centers can still create longer-term economic benefits through indirect channels. While the regression only directly measures the relationship between data center treatment and log GDP, the existing literature suggests several possible mechanisms behind this relationship. PwC (2024) emphasizes construction, operations, and indirect local spending; Muro, Parilla, and Maxim (2026) highlight tax revenue, infrastructure investment, and limited employment effects; and Park (2025) suggests that proximity to data centers may improve firm productivity. These studies collectively point toward data centers influencing GDP through indirect channels rather than serving as the sole driver of growth. They may increase local tax revenue, attract related technology firms, improve digital infrastructure, and support businesses that depend on cloud computing and low-latency data services.

The results also highlight an important tradeoff for local governments. Many cities and counties offer tax incentives to attract data centers, but these incentives can reduce the financial benefits that communities expect to receive (Muro, Parilla, & Maxim, 2026). If a county gives up too much tax revenue in exchange for a project that creates few permanent jobs, the long-term economic payoff may be weaker than advertised. In addition, data centers place pressure on local energy and water systems, which can create costs that are not fully captured by GDP growth alone (Ngata et al., 2025).

Overall, the findings suggest that data centers are associated with local economic growth, but the relationship is more complex than simply just that data centers cause growth. The regression results show a statistically significant positive association, but this relationship is best understood as indirect and conditional rather than as data centers being the only cause of higher GDP. Their benefits appear strongest when they are connected to broader regional development like technology and infrastructure improvements. In areas where data centers operate mostly as stand-alone facilities, the effects may be more short-lived and limited. Therefore, the evidence supports a mixed conclusion: data centers can contribute to sustained local economic growth, but mainly when local communities are able to convert the initial investment into broader economic development.

CONCLUSION

This paper examined whether counties with major data centers show higher GDP, and what that pattern suggests about the longer-run question of persistence. Based on the regression results, data center treatment is positively and significantly associated with higher county-level GDP. The two-way fixed effects model found a positive and statistically significant coefficient of 0.1609, which translates to an estimated 17.46% increase in GDP for treated counties, holding county-specific and year-specific factors constant. This suggests that data centers are not just minor additions to local economies; they can be connected to meaningful increases in economic output.That said, the model averages across all post-treatment years and cannot distinguish effects that hold over time from those that fade after the initial investment push. The findings are consistent with sustained growth, but do not directly demonstrate it.

The results should also be interpreted carefully for other reasons. Counties that attract data centers may already have advantages such as strong infrastructure, access to electricity, available land, business-friendly policies, or proximity to growing metropolitan areas. In that sense, data centers may be part of a larger growth pattern rather than the only reason for it.

The findings support a mixed conclusion. Data centers can help local economies, especially when they encourage infrastructure investment, business activity, technology clustering, and long-term regional development. But their benefits are weaker when they operate as stand-alone facilities with few connections to the surrounding economy. Since data centers are capital intensive and do not always create many permanent jobs, the strongest direct effects may occur during the construction and investment phase. Long-term benefits depend on whether communities can turn that initial investment into broader economic gains.

Data centers can contribute to local economic growth, but the size and duration of that contribution depend heavily on local conditions. For local governments, the key issue is not just attracting a data center, but making sure the project creates lasting value for communities in the long run. Future research should use a larger sample of counties, include more detailed data on facility size, and account for additional local factors such as employment, population growth, tax incentives, and energy costs. Future work should also use event study or dynamic treatment effect designs, which would show whether the GDP association strengthens, holds, or erodes as more years pass since a data center opened. Doing so would give a clearer picture of when data centers truly support long-term development and when their effects remain short-lived.

REFERENCES

Chao Li, Wanling He, Erbao Cao. (2025). *Impact of green data center pilots on digital economy development: Evidence from Chinese cities*.
https://dl.acm.org/doi/10.1016/j.cie.2025.110914

Muro, M., Parilla, J., & Maxim, R. (2026). *The local implications of data centers for rural communities in the US*. Brookings Institution.
https://www.brookings.edu/articles/local-implications-data-centers-rural-communities-us/

Wacuka Ngata, Noman Bashir, Michelle Westerlaken, Laurent Liote, Yasra Chandio, Elsa Olivetti. (2025). *The cloud next door: Investigating the environmental and socioeconomic strain of datacenters on local communities*. arXiv.
https://arxiv.org/abs/2506.03367

Jiyong Park. (2025). *Data (center) economy: The impact of data centers on firm performance*. In *Proceedings of the Americas Conference on Information Systems (AMCIS 2025)*.
https://aisel.aisnet.org/amcis2025/ict_global/ict_global/5/

PricewaterhouseCoopers (PwC). (2024). *Economic, environmental, and social impacts of data centers in the United States*. PwC.
https://psc.ky.gov/pscecf/2024-00326/rick.lovekamp%40lge-ku.com/12182024124733/07-PSC_DR1_LGE_KU_Attach_to_Q20%28a%29_-_Data_Center_Impact_Study.pdf