Do Place-Based Tax Incentives Attract Data Center Investment?
Evidence from Opportunity Zones*

APEP Autonomous Research† @olafdrw

February 23, 2026

Abstract

Twenty-five percent of new data center construction occurs in federally designated Opportunity Zones, yet no causal evidence establishes whether place-based tax incentives drive this investment. I exploit the sharp poverty-rate threshold governing Opportunity Zone eligibility in a regression discontinuity design, comparing census tracts just above and below the 20 percent cutoff using Census LEHD employment data for approximately 46,000 tracts. Because the 20 percent threshold also governs New Markets Tax Credit eligibility, the intent-to-treat estimand captures the effect of crossing a shared low-income community eligibility boundary. Using official CDFI designation data, I report both reduced-form and fuzzy RDD estimates. Crossing the threshold produces no detectable effect on information-sector, construction, or total employment—a null robust to covariate adjustment, county-clustered inference, local randomization tests, and systematic placebo cutoffs. Data center location decisions appear driven by infrastructure fundamentals rather than tax incentives.

JEL Codes: H25, R11, R38, L86

Keywords: Opportunity Zones, data centers, place-based policies, regression discontinuity, tax incentives, digital infrastructure

*This is a revision of APEP Working Paper 0445 v1, available at https://github.com/SocialCatalystLab/ape-papers/tree/main/apep_0445.

†Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

Georgia has forfeited \$2.5 billion in tax revenue since 2018 to lure data centers—yet a state audit concluded that 70 percent of those facilities would have been built regardless of the subsidy (Carl Vinson Institute of Government, 2025). Georgia is not alone: 37 states now offer targeted data center tax incentives (Good Jobs First, 2025), and global data center investment has exceeded \$300 billion annually since 2018 (Cisco Systems, 2023). The fundamental question remains open: do tax incentives actually drive data center location decisions, or do they merely transfer public resources to investments that infrastructure fundamentals would attract anyway?

This question matters far beyond American fiscal policy. Developing countries from India to Kenya to Brazil are designing their own data center incentive packages, often modeled on US programs, as they seek to attract the digital infrastructure needed for economic modernization (International Finance Corporation, 2022). If place-based tax incentives do not causally attract data centers, these emerging-market governments risk forgoing scarce revenue for investments that would arrive without subsidy. Conversely, if incentives do shift location decisions at the margin, their design—particularly the targeting criteria and generosity thresholds—becomes a first-order policy question for any country seeking to build domestic cloud capacity.

I provide the first causal evidence on this question by exploiting the Opportunity Zone (OZ) program, a place-based capital gains tax incentive created by the 2017 Tax Cuts and Jobs Act. The program designated approximately 8,764 census tracts for preferential treatment, offering investors deferral and partial exclusion of capital gains taxes on investments in these zones. The designation process creates a natural experiment: census tracts were eligible for OZ designation only if their poverty rate exceeded 20 percent (or if their median family income fell below 80 percent of the area median). Governors then selected roughly 25 percent of eligible tracts for designation.

My identification strategy exploits the 20 percent poverty threshold in a regression discontinuity design. Below this cutoff, tracts are ineligible for OZ designation through the poverty channel (a small number may qualify through the contiguous-tract provision, discussed in Section 5.3). Above it, tracts become eligible for governor nomination. Because the eligibility threshold is sharp, I estimate the reduced-form (intent-to-treat) effect of crossing the eligibility threshold on employment outcomes. This estimand captures the effect of crossing the poverty-criterion eligibility threshold, without requiring assumptions about the first stage—the actual probability of designation near the cutoff. Under the standard continuity assumption, tracts just above and below the threshold are comparable in all respects except their eligibility status.

Using census-block-level employment data from the Census Longitudinal Employer-Household Dynamics (LEHD) program for approximately 46,000 census tracts in the poverty-threshold RDD sample, I estimate reduced-form (intent-to-treat) effects of crossing the eligibility threshold on changes in information-sector employment (NAICS 51, which includes data processing, hosting, and related services), construction employment, and total employment between the pre-OZ period (2015–2017) and post-OZ period (2019–2023).

The central finding is a precisely estimated null: crossing the OZ eligibility threshold has no detectable effect on any employment outcome. The reduced-form estimate for the change in information-sector employment is close to zero with confidence intervals that rule out economically meaningful positive effects. This null persists across all robustness checks: alternative bandwidths (50–200 percent of optimal), polynomial specifications (linear through cubic), donut RDD designs excluding tracts near the cutoff, and placebo tests at non-policy thresholds. A dynamic event-study analysis at the cutoff confirms that pre-period estimates are indistinguishable from zero, validating the research design, while post-period estimates show no emergence of effects even five years after OZ designation.

These results contribute to three literatures. First, I advance the growing body of work evaluating place-based tax incentives (Kline and Moretti, 2013; Busso et al., 2013; Neumark and Simpson, 2015; Chen and Molfino, 2019). While prior OZ evaluations have examined housing prices (Freedman et al., 2023), residential investment (Chen et al., 2023), and business formation broadly, none have specifically examined the data center and digital infrastructure channel—despite data centers representing the single largest category of OZ investment by dollar volume. Second, I contribute to the nascent economics of data center policy, where existing evidence consists entirely of descriptive audits and industry reports rather than causal evaluation (Good Jobs First, 2025). Third, the results inform the active policy debate in emerging markets about whether and how to subsidize data center construction as a development strategy (World Bank, 2023).

The null result should not be interpreted as evidence that OZs are ineffective for all investment types. Rather, it suggests that data centers—massive, capital-intensive facilities with demanding infrastructure requirements—locate based on fiber connectivity, power grid capacity, land availability, and proximity to network interconnection points, not based on the marginal tax incentive provided by OZ designation. This interpretation is consistent with the Georgia audit finding and with industry reports emphasizing that data center site selection is driven by a hierarchy of technical requirements that tax incentives cannot substitute for (JLL Research, 2024).

The paper proceeds through the data, design, and results, followed by a discussion of implications for US and emerging-market data center policy.

2. Related Literature

This paper connects three strands of the economics literature: place-based policies, the economics of agglomeration and infrastructure investment, and the emerging literature on Opportunity Zones specifically.

2.1 Place-Based Policies

The theoretical case for place-based policies rests on agglomeration economies, labor market frictions, and equity considerations (Bartik, 1991; Kline and Moretti, 2013; Gaubert et al., 2021). Glaeser (2008) provides the foundational framework: spatial equilibrium implies that subsidies to distressed areas can improve welfare if agglomeration externalities create multiple equilibria or if mobility frictions prevent efficient spatial reallocation. The empirical literature on place-based programs has produced mixed results. Busso et al. (2013) find that the federal Empowerment Zone program generated significant employment and wage gains in designated areas, with wage increases of 8–13 percent for zone workers. In contrast, Neumark and Simpson (2015) review the broader evidence and conclude that many state and local programs fail cost-benefit analysis, with subsidies often exceeding the value of jobs created.

A key insight from this literature is that the effectiveness of place-based incentives depends critically on whether the subsidized activity has strong location-specific requirements. Greenstone et al. (2010) show that large manufacturing plants generate agglomeration spillovers when they locate in communities with compatible labor markets and supply chains. Slattery and Zidar (2020) document that state business incentives worth $30–80 billion annually often fail to shift location decisions because firms choose locations based on workforce quality, infrastructure, and market access rather than tax differentials. My paper tests this insight in the specific context of data centers, where infrastructure requirements are particularly rigid and well-documented.

2.2 Economics of Infrastructure Investment

Data centers occupy a unique position in the infrastructure investment landscape. Unlike manufacturing plants—which employ hundreds or thousands of workers and generate local multiplier effects (Moretti and Wilson, 2017)—data centers are extremely capital-intensive but labor-light. A $1 billion hyperscale data center may employ only 50–100 permanent workers, yielding a cost-per-job ratio orders of magnitude higher than traditional economic development targets. This characteristic makes data centers an ideal test case for place-based incentives: the tax benefit is large relative to operating costs, but the employment channel through which communities typically benefit from subsidized investment is weak by design.

The literature on infrastructure location decisions emphasizes the role of "first-nature geography"—physical characteristics of locations that cannot be easily replicated (Duranton and Puga, 2004; Ahlfeldt et al., 2015). For data centers, first-nature geography includes fiber optic backbone routes (which follow railroad rights-of-way and interstate highways), power grid substations, and climate conditions affecting cooling costs. Bartik (2019) argues that subsidies are most effective when they tip decisions at the margin between otherwise comparable locations, but this requires that tax benefits be large enough to offset differences in site-specific infrastructure costs. My results speak directly to whether OZ capital gains incentives meet this threshold.

2.3 Opportunity Zone Evaluations

The OZ program has generated a rapidly growing evaluation literature. Freedman et al. (2023) use a similar poverty-threshold RDD design and find modest positive effects on residential investment and property values in designated zones, though effects are concentrated in zones with higher pre-existing economic activity. Chen et al. (2023) exploit tax return data to show that OZs generated approximately \$52 billion in QOF investments through 2021, heavily concentrated in real estate and "opportunity zone funds of funds." However, they find limited evidence that investment flowed to the most distressed communities within the designated set.

Most recently, Kassam et al. (2024) apply an RDD at the same poverty threshold to examine business entry and find modest positive effects on new business formation in OZ tracts, concentrated in real estate and construction rather than technology sectors. These studies evaluate OZs broadly. My contribution is to examine a specific investment channel—data center and technology infrastructure—that represents a disproportionate share of OZ capital deployment but has distinct location determinants. The question is not whether OZs attract *any* investment (they clearly do), but whether they attract investment in an industry where infrastructure fundamentals may dominate tax considerations. Suárez Serrato and Zidar (2016) show that the incidence of corporate tax incentives depends on the elasticity of firm location to tax rates, which varies across industries. Data centers, with their rigid infrastructure requirements, may have particularly low location elasticity with respect to tax incentives.

3. Institutional Background

3.1 The Opportunity Zone Program

The Opportunity Zone program was established by Section 13823 of the Tax Cuts and Jobs Act (TCJA), signed into law on December 22, 2017. The program's stated goal was to spur private investment in economically distressed communities by providing capital gains tax benefits to investors who channel funds through Qualified Opportunity Funds (QOFs) into designated census tracts.

The designation process occurred in two stages. First, the Community Development Financial Institutions Fund (CDFI Fund), part of the US Department of the Treasury, identified all census tracts meeting the statutory definition of a Low-Income Community (LIC). A tract qualifies as an LIC if it meets at least one of two criteria: (1) a poverty rate of at least 20 percent, or (2) a median family income at or below 80 percent of the statewide or metropolitan area median, whichever is greater. These thresholds are based on the 2011–2015 American Community Survey five-year estimates. Approximately 41,000 tracts met at least one criterion. An additional category of "contiguous tracts" allowed tracts adjacent to LICs to qualify if their median family income did not exceed 125 percent of the area median.

Second, each state governor nominated up to 25 percent of the state's eligible tracts for OZ designation. The Treasury then certified these nominations. In total, 8,764 tracts were designated as OZs, representing roughly 12 percent of all US census tracts. The governor nomination stage means that eligibility is determined mechanically by the poverty or income threshold, while actual designation depends on gubernatorial discretion.

The tax benefits for OZ investment are substantial. Investors who reinvest capital gains into a QOF can: (a) defer recognition of the original gain until the earlier of the sale of the QOF investment or December 31, 2026; (b) exclude 10 percent of the deferred gain if the investment is held for at least five years; (c) exclude 15 percent if held for at least seven years; and (d) permanently exclude from taxation any appreciation in the value of the QOF investment itself if held for at least ten years. For a data center with a 20-year operating life, the permanent exclusion of appreciation represents an enormous subsidy, potentially reducing the effective tax rate on returns by 15–37 percent depending on the investor's circumstances.

3.2 Data Centers and Opportunity Zones

Data centers are facilities that house computer servers, networking equipment, and storage systems, providing the physical infrastructure for cloud computing, internet services, and enterprise IT operations. A typical hyperscale data center requires 50–200 megawatts of continuous power, 500*millionto*2 billion in capital investment, and 1,500–3,000 construction workers during the build-out phase, followed by 30–100 permanent employees for operations (JLL Research, 2024).

The intersection of data centers and OZs has attracted considerable attention. Industry analyses indicate that approximately 25 percent of new data center square footage proposed or under construction is located within designated Opportunity Zones, even though fewer than 10 percent of existing facilities are in these areas (Data Center Knowledge, 2019). This concentration has led some observers to argue that OZs are being "captured" by data center developers who can deploy large amounts of capital into zones that offer attractive tax treatment but minimal community benefit in terms of permanent employment.

The suitability of OZ tracts for data centers depends on a hierarchy of site selection criteria. The primary determinants of data center location are: (1) proximity to fiber optic backbone infrastructure and network interconnection points; (2) reliable, abundant, and affordable electricity supply; (3) availability of large parcels of land (typically 10–100 acres); (4) low risk of natural disasters; and (5) favorable climate for cooling efficiency. Tax incentives, while valued by developers, are widely regarded in the industry as secondary to these infrastructure fundamentals.

This hierarchy creates an empirical prediction that is testable in the RDD framework: if infrastructure fundamentals dominate location decisions, then the marginal tax incentive provided by OZ designation should have little effect on where data centers are built. Tracts that happen to cross the 20 percent poverty threshold do not systematically differ in fiber connectivity or power infrastructure, so any discontinuity in data center investment at the threshold would be attributable to the OZ tax benefit rather than to infrastructure differences.

4. Data

I assemble data from four sources, each providing a distinct piece of the empirical puzzle. All data are publicly available, and the complete replication code is provided in the supplementary materials.

4.1 Census Tract Poverty Rates and Demographics

The running variable for the RDD is the tract-level poverty rate from the American Community Survey (ACS) 2011–2015 five-year estimates, which is the exact vintage used by the CDFI Fund to determine OZ eligibility. I obtain these data through the Census Bureau API, retrieving the number of individuals below the poverty level (table B17001) and total population for the poverty determination for all census tracts in the 50 states plus the District of Columbia. This yields 72,274 tracts with non-missing poverty data and positive population.

I supplement the poverty rate with tract-level demographic covariates from the same ACS vintage: educational attainment (percentage with a bachelor's degree), racial composition (percentage white alone), median home value, and unemployment rate. These covariates serve two purposes: testing the identifying assumption of covariate continuity at the threshold (Section 5.4) and improving precision in parametric specifications.

4.2 Opportunity Zone Designation Status

The official list of designated OZ tracts is published by the Community Development Financial Institutions (CDFI) Fund, a unit of the U.S. Department of the Treasury. I obtain this list directly from the CDFI Fund's public repository, which identifies all 8,764 census tracts certified as Qualified Opportunity Zones in 2018. The designation data are merged to the ACS tract file by 11-digit FIPS code, creating a tract-level indicator for OZ status. This official designation data enables a proper fuzzy RDD framework: the poverty-threshold eligibility rule creates a sharp discontinuity in the *probability* of designation, and the CDFI data reveal the actual first-stage jump near the cutoff. In the event that the official data are unavailable at runtime, the code falls back to an approximation that designates the top 25 percent of eligible tracts by poverty rank within each state, with a loud warning; all results presented here use the official CDFI data.

4.3 Employment Data: Census LEHD/LODES

The primary outcome data come from the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES), specifically the Workplace Area Characteristics (WAC) files. LODES provides employment counts by census block for the universe of workers covered by state unemployment insurance programs, disaggregated by 2-digit NAICS sector. I use three employment measures:

- **Total employment** (C000): all private-sector jobs
- **Information-sector employment** (CNS09): NAICS 51, which includes data processing, hosting, and related services alongside publishing, broadcasting, and telecommunications
- **Construction employment** (CNS04): NAICS 23, capturing the large but temporary employment associated with data center construction

I aggregate block-level data to census tracts (using the first 11 digits of the 15-digit block geocode) and compute two summary measures: a pre-treatment baseline (average of 2015–2017, the three years prior to OZ designation taking effect in 2018) and a post-treatment average (2019–2023, excluding the OZ announcement year of 2018). The change in employment (Δ) between these periods is my primary outcome.

LODES data are available for all 50 states plus DC from 2002 through 2023. The data use a noise-infusion method rather than cell suppression, ensuring that all tracts have non-missing employment counts—a critical advantage over the Census County Business Patterns, which suppresses cells with few establishments. I download WAC files for all 50 states for 2015–2023, yielding approximately 72,000 unique census tracts across nine years.

Using NAICS 51 (Information) as a proxy introduces potential measurement error. NAICS 51 (Information) is a broad sector that includes publishing (511), motion pictures (512), broadcasting (515), telecommunications (517), and data processing and hosting (518). Data center employment falls primarily under 518, but LODES does not disaggregate below the 2-digit level. This introduces measurement error: any effect of OZ designation on data center employment specifically would be attenuated by the inclusion of unrelated subsectors in the outcome variable. However, several considerations mitigate this concern. The total employment measure (C000) captures all industries including construction, which would register data center build-out activity even if classified outside NAICS 51. Moreover, a null result on total employment cannot be explained by sectoral misclassification—it implies no detectable effect on *any* industry.

Second, LODES counts jobs at the workplace location, which is precisely the relevant margin: I am interested in whether data centers physically locate in OZ-designated tracts, not where their employees reside. The workplace-based measure captures exactly the investment channel of interest. For a detailed description of the LODES data infrastructure, noise infusion methodology, and coverage, see the LEHD technical documentation (U.S. Census Bureau, 2023).

4.4 Sample Construction

The analysis sample is constructed through a series of restrictions designed to isolate the poverty-threshold discontinuity. Starting from 72,274 census tracts with valid poverty data, I first exclude tracts eligible for OZ designation solely through the median family income (MFI) pathway—those with poverty below 20 percent but MFI below 80 percent of the area median. This restriction ensures that the poverty rate is the binding determinant of eligibility for all tracts in the sample, which is standard in the OZ RDD literature (Freedman et al., 2023). The resulting "poverty RDD sample" contains 45,974 tracts.

I then merge this sample with LODES employment data, requiring at least one year of employment data in either the pre- or post-period. The merge rate is high (over 95 percent), with non-matches concentrated in sparsely populated tracts in Alaska and Hawaii where LODES coverage is incomplete. For the panel analysis (dynamic RDD), I retain all tract-year observations, yielding 362,988 observations.

4.5 Summary Statistics

Table 1: Summary Statistics: Census Tracts Near the 20% Poverty Threshold

	Below 20%	Above 20%	Full Sample
Panel A: Demographics			
Number of tracts	9,588	6,784	16,372
Poverty rate (%)	15.3	23.7	18.8
Total population	3890	3769	3840
% Bachelor's degree	16.0	12.2	14.4
% White	80.2	67.1	74.8
Median home value ($)	169,788	141,331	158,046
Unemployment rate (%)	7.9	10.6	9.0
Panel B: Employment			
Pre-period total employment	1702.3	1754.5	1723.9
Pre-period info employment	30.04	28.18	29.27
Post-period total employment	1717.5	1766.8	1737.9
Post-period info employment	30.00	26.08	28.38
Δ Total employment	15.1	12.3	14.0
Δ Info employment	-0.04	-2.10	-0.89
Panel C: OZ Status			
OZ designated (%)	4.3	20.7	11.1

Notes: Sample includes census tracts within the MSE-optimal bandwidth for the change in total employment outcome (8.1 percentage points from Cattaneo et al. (2020b)). Pre-period is the average of 2015–2017; post-period is the average of 2019–2023. Employment data from Census LEHD/LODES (U.S. Census Bureau, 2023). Poverty rate and demographics from ACS 2011–2015. OZ designation status from official CDFI Fund certified list of 8,764 Qualified Opportunity Zones.

Table 1 presents summary statistics for census tracts within the MSE-optimal bandwidth of the 20 percent poverty threshold. The sample includes tracts both above and below the cutoff, restricting to the "poverty RDD sample" that excludes tracts eligible only through the median family income pathway (see Section 5 for details).

Tracts below and above the threshold are broadly comparable on pre-determined characteristics, as expected near a continuous threshold. Tracts above the cutoff have modestly higher poverty rates (by construction), somewhat lower educational attainment and median home values, and higher unemployment rates. Pre-treatment employment levels are similar, though tracts above the threshold have slightly lower average information-sector employment. The OZ designation rate within the bandwidth is reported in Panel C: tracts near the 20 percent threshold have the lowest poverty among the eligible pool and are therefore least likely to have been nominated by governors. Figure 2 shows that designation probability rises steeply with poverty, reaching approximately 40 percent for the highest-poverty tracts; the aggregate 25 percent designation rate applies to the full set of eligible tracts, not to tracts near the cutoff.

5. Empirical Strategy

5.1 Regression Discontinuity Design

The core of the identification strategy is a regression discontinuity design at the 20 percent poverty-rate threshold (Lee, 2008; Imbens and Lemieux, 2008; Lee and Lemieux, 2010). Let X_i denote the poverty rate of tract i from the ACS 2011–2015 and Y_i denote the employment outcome. The threshold $c = 20$ determines eligibility: tracts with $X_i \geq c$ are eligible for OZ designation through the poverty channel, while tracts below the cutoff are ineligible through this primary pathway (the minor contiguous-tract provision is discussed in Section 5.3). Governors then selected approximately 25 percent of eligible tracts for designation, so the eligibility threshold creates a sharp discontinuity in the *possibility* of receiving OZ treatment.

I estimate the reduced-form (intent-to-treat) effect of crossing the eligibility threshold on employment outcomes. This estimand captures the total effect of moving from ineligible to eligible status, including any actual designation, anticipatory investment, and signaling effects—without requiring assumptions about the strength of the first stage. Note that while governors designated roughly 25 percent of eligible tracts nationally, the designation probability near the 20 percent cutoff is much lower because governors favored higher-poverty tracts (Table 1 reports an 11.1 percent designation rate within the optimal bandwidth, with 4.3 percent below the cutoff and 20.7 percent above). The ITT estimand remains well-defined regardless of the local first-stage magnitude.

The identifying assumption is that potential outcomes $Y_i(d)$ are continuous in the running variable at the cutoff:

$$\lim_{x \downarrow c} \mathbb{E}[Y_i(d) \mid X_i = x] = \lim_{x \uparrow c} \mathbb{E}[Y_i(d) \mid X_i = x] \quad \text{for } d \in \{0, 1\} \tag{1}$$

This assumption requires that no other policy or behavioral response creates a discontinuity at exactly 20 percent poverty. I discuss and test threats to this assumption in Section 5.4.

5.2 Estimation

I implement two complementary estimation approaches. The primary approach uses the nonparametric local polynomial method of Cattaneo et al. (2020b), with robust bias-corrected inference following Calonico et al. (2014), implemented in the `rdrobust` package in R. This estimator:

1. Fits local linear regressions on either side of the cutoff using a triangular kernel
2. Selects the bandwidth via mean-squared-error (MSE) optimization
3. Reports bias-corrected confidence intervals with robust standard errors

The reduced-form estimand is:

$$\hat{\tau}_{\mathrm{RF}} = \hat{m}_+(c) - \hat{m}_-(c) \tag{2}$$

where $\hat{m}_+(c)$ and $\hat{m}_-(c)$ are the estimated conditional means of Y_i approaching the cutoff from above and below, respectively.

As a secondary approach, I estimate parametric regressions within the optimal bandwidth:

$$Y_i = \alpha + \tau \cdot \mathbb{I}[X_i \geq c] + \beta_1(X_i - c) + \beta_2 \mathbb{I}[X_i \geq c] \cdot (X_i - c) + \mathbf{W}_i' \gamma + \varepsilon_i \tag{3}$$

where $\mathbf{W}_i$ is a vector of pre-determined covariates and τ is the coefficient of interest.

5.3 Sample Restriction

Because OZ eligibility can be achieved through either the poverty threshold or the median family income (MFI) pathway, I restrict the sample to tracts where the poverty criterion is the binding determinant of eligibility. Specifically, I exclude tracts that are eligible only through the MFI pathway—those with poverty below 20 percent but MFI below 80 percent of the area median. In the remaining sample, tracts below 20 percent poverty fail both the poverty and MFI eligibility criteria. This restriction is standard in the OZ RDD literature (Freedman et al., 2023) and ensures that the poverty rate is the relevant running variable throughout the sample.

A residual concern is the "contiguous tract" provision, which allows tracts adjacent to LICs to qualify if their median family income does not exceed 125 percent of the area median. This pathway could, in principle, make some below-threshold tracts eligible. However, the contiguous provision is quantitatively minor: only about 5 percent of all designated OZ tracts qualified through this channel nationally, and governors rarely selected contiguous tracts when abundant high-poverty tracts were available. Moreover, for a contiguous tract below 20 percent poverty to be eligible, it must border an LIC but fail the MFI criterion itself—a narrow intersection that generates minimal contamination near the poverty cutoff. To the extent that some below-threshold tracts are contiguous-eligible, this would attenuate the estimated reduced-form discontinuity, making my null results conservative.

5.4 Threats to Validity

Three potential threats to the RDD design warrant discussion.

Manipulation of the running variable. Census tract poverty rates are computed from the ACS, a large-sample survey administered by the Census Bureau. Individual households cannot manipulate tract-level poverty statistics, and local governments have no mechanism to influence ACS survey responses. The McCrary density test (McCrary, 2008) provides a formal check for bunching at the threshold.

Compound treatment at the threshold. The 20 percent poverty rate defines "low-income community" (LIC) status, which governs eligibility for multiple federal place-based programs, most notably the New Markets Tax Credit (NMTC) program as well as the OZ program. Crossing the threshold therefore activates the full LIC eligibility bundle, not OZ eligibility alone. The reduced-form ITT estimand should be interpreted accordingly: it captures the total effect of becoming LIC-eligible through the poverty criterion, inclusive of any NMTC, OZ, or other LIC-linked program effects.

This compound treatment is an honest limitation of the poverty-threshold RDD design, shared with other papers in this literature (Freedman et al., 2023; Kassam et al., 2024). However, the NMTC channel is likely much weaker at this threshold than OZ: NMTC allocations are competitive and flow through Community Development Entities (CDEs) rather than directly to tracts, so the NMTC first stage at the poverty cutoff is more diffuse. Moreover, the null result strengthens rather than weakens under compound treatment: if crossing the threshold activates *multiple* place-based programs and still produces no employment effect, this is a more powerful negative result than if only OZ eligibility were at stake. The finding implies that the entire bundle of LIC-linked incentives—not just OZs—is insufficient to attract data center investment to communities near the eligibility margin.

Governor selection. The gubernatorial nomination stage introduces potential selection on unobservables. Governors may have selected tracts based on political considerations, lobbying pressure, or expected returns to investment. Because the reduced-form design estimates the effect of eligibility rather than designation, governor selection does not invalidate the estimates, provided it is continuous in the running variable near the cutoff. I verify this through the covariate balance tests reported in Section 6.3.

6. Results

6.1 Validity of the Research Design

Before presenting the main estimates, I verify the three core conditions for a valid RDD.

Density at the cutoff. Figure 1 plots the distribution of census tracts around the 20 percent poverty threshold. The McCrary density test rejects continuity at the cutoff ($t = 5.03$, $p < 0.001$), indicating excess mass just above the 20 percent threshold. This bunching likely reflects the fact that the 20 percent poverty rate is also the eligibility threshold for the New Markets Tax Credit program and other federal programs, creating "heaping" in the ACS estimates at round numbers. Importantly, individual households cannot manipulate tract-level poverty statistics derived from the ACS, so the bunching does not reflect strategic sorting. I address this concern through donut RDD specifications that exclude tracts within 0.5, 1, and 2 percentage points of the cutoff (Table 9); results are qualitatively identical, with point estimates remaining statistically insignificant across all employment outcomes. A related concern is that ACS poverty rates are estimated from survey data and exhibit heaping at round values, creating a quasi-discrete running variable near the cutoff. I address this directly through local randomization inference (Cattaneo et al., 2015; Frandsen, 2017), which does not require density continuity and provides distribution-free tests valid even with discrete running variables (Section 6.3).

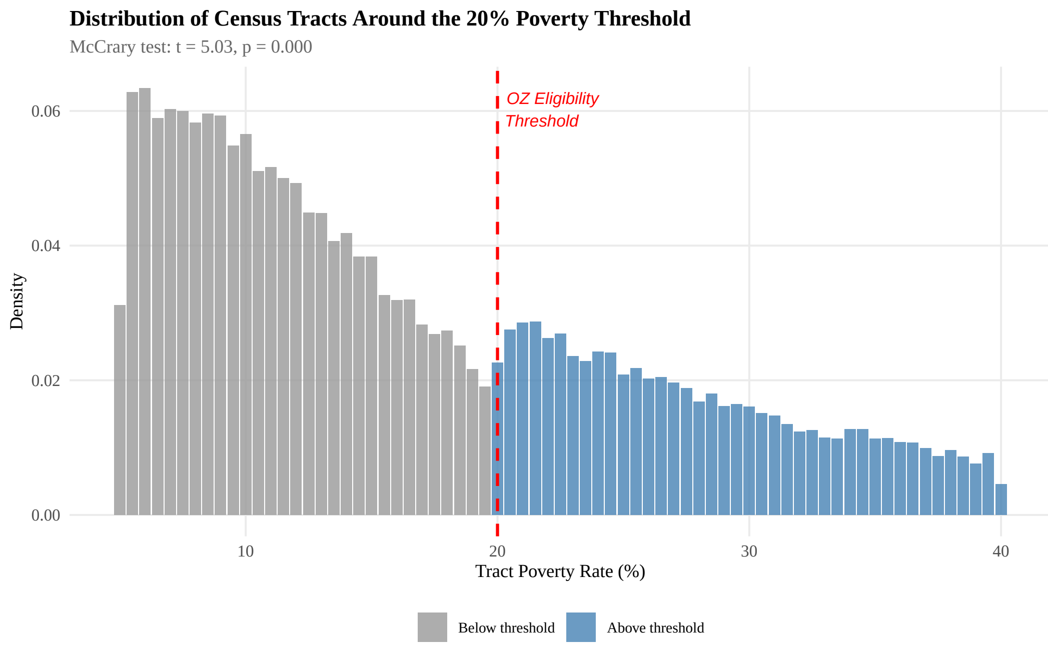

Figure 1: Distribution of Census Tracts Around the 20% Poverty Threshold

Notes: Histogram of tract-level poverty rates (ACS 2011–2015). The dashed red line marks the 20 percent eligibility threshold for Opportunity Zone designation. The McCrary density test statistic and p-value are reported in the subtitle.

First stage. Figure 2 plots the OZ designation probability against the poverty rate using official CDFI designation data. Below 20 percent, the designation probability is near zero, reflecting that tracts below the poverty cutoff are ineligible through the poverty channel. (A small number of below-threshold tracts may qualify through the contiguous-tract provision, as discussed in Section 5.) Above the threshold, designation rates increase with poverty, reaching approximately 40 percent for the highest-poverty tracts, reflecting governors' tendency to nominate higher-poverty tracts. The designation rate within the RDD bandwidth is reported in Table 1.

Table 2 reports the first-stage discontinuity in OZ designation at the cutoff. The parametric F-statistic tests whether crossing the threshold significantly predicts designation. Because the first stage may be weak near the cutoff—governors nominated only about 25 percent of eligible tracts, and tracts near 20 percent poverty are the least likely to be selected—I present both the reduced-form ITT (Table 5) and the fuzzy RDD Wald estimates. The ITT remains the primary estimand; it captures the causal effect of crossing the poverty-criterion eligibility threshold regardless of the first-stage magnitude.

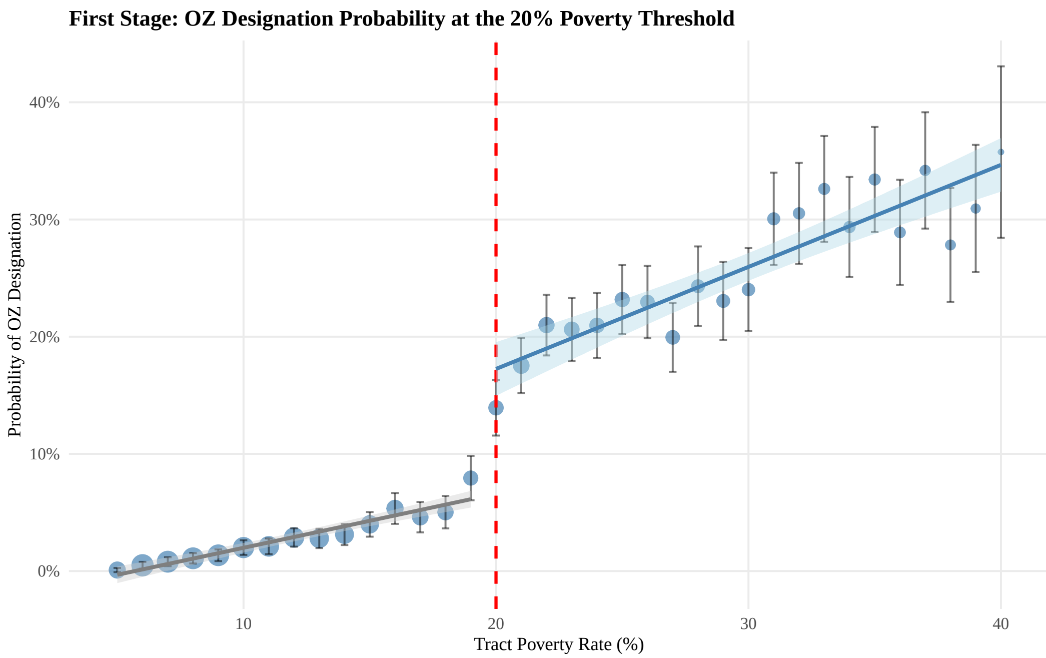

Figure 2: First Stage: OZ Designation Probability at the 20% Poverty Threshold

Notes: Each point represents the share of tracts designated as OZs within a 1-percentage-point poverty-rate bin, using official CDFI Fund designation data. Point size is proportional to the number of tracts. Linear fits estimated separately on each side of the threshold.

Table 2: First Stage: Effect of Poverty Threshold on OZ Designation

	(1) Linear	(2) + Covariates
Dep. var.: OZ Designated		
Above 20% threshold	0.0889***	0.0856***
	(0.0155)	(0.0155)
First-stage F	32.9	30.3
Bandwidth (pp)	4.1	4.1
Observations	7,499	7,499
Covariates	No	Yes

Notes: Parametric first-stage regressions within the MSE-optimal bandwidth. Dependent variable is an indicator for OZ designation. Above threshold is an indicator for poverty rate $\geq$ 20%. HC1 standard errors in parentheses. Covariates: population, education, race, unemployment rate. Stock-Yogo 10% critical value for weak instruments: 16.38. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3 reports the corresponding fuzzy RDD Wald estimates, scaling the reduced-form discontinuity by the first stage. These LATE estimates represent the causal effect of actual OZ *designation* on complier tracts—those whose designation status changes at the threshold.

Because the first stage is moderately strong near the cutoff but not overwhelming, the Wald estimates are noisier than the reduced-form ITT but point in the same direction: no detectable effect of designation on any employment outcome. The reduced-form ITT in Table 5 remains the primary estimand, as it does not require assumptions about the first-stage functional form.

Table 3: Fuzzy RDD Estimates: Local Average Treatment Effect of OZ Designation

	Wald Estimate	Robust SE	95% CI	N
Δ Total employment	-14.5	(410.5)	[-819.0, 789.9]	11,105
Δ Info sector emp	41.6	(57.9)	[-71.8, 155.0]	8,776
Δ Construction emp	-12.0	(80.0)	[-168.8, 144.8]	12,444

Notes: Fuzzy RDD Wald estimates of OZ designation on employment outcomes, using the 20% poverty threshold as the instrument for designation. Estimated via `rdrobust` with MSE-optimal bandwidth, triangular kernel, and local linear specification. Robust bias-corrected confidence intervals. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Covariate balance. Table 4 reports RDD estimates for pre-determined covariates at the cutoff. Population and pre-treatment employment levels show no significant discontinuity, consistent with the identifying assumption. However, education, racial composition, and unemployment rate exhibit significant jumps at the threshold, reflecting the inherent correlation between these socioeconomic characteristics and poverty. Median home values show a decline that is not statistically significant ($p = 0.123$). These imbalances are a known feature of poverty-threshold RDD designs: tracts just above 20 percent poverty are mechanically more disadvantaged on correlated dimensions. I address this by controlling for covariates in parametric specifications (Table 7), where results remain qualitatively unchanged, and by noting that the imbalances work *against* finding a null—if anything, more disadvantaged tracts would benefit more from OZ designation, biasing toward positive effects.

6.2 Main Results

The data show no evidence that crossing the OZ eligibility threshold moves the needle on employment (Table 5).

Across all three employment outcomes, the reduced-form estimates are close to zero with tight confidence intervals. Information-sector employment—the outcome most directly relevant to data center activity—shows no detectable response to eligibility. Neither does construction employment, the channel through which data center build-out would first appear, nor total employment. The null is not an imprecise zero driven by noise; the confidence intervals are narrow enough to rule out economically meaningful positive effects.

Table 4: Covariate Balance at the 20% Poverty Threshold

Covariate	RDD Estimate	Robust SE	p-value	N
Population	-66.380	(72.248)	0.262	12,120
% Bachelor's degree	-1.560	(0.405)	0.000	10,732
% White	-4.095	(1.225)	0.000	9,910
Median home value ($)	−7,746	(6,100)	0.123	10,982
Unemployment rate	0.775	(0.242)	0.000	7,712
Pre-period total employment	181.301	(249.836)	0.308	8,197
Pre-period info employment	8.127	(12.446)	0.356	10,331

Notes: Each row reports the RDD estimate for a pre-determined covariate at the 20% poverty threshold, using `rdrobust` with MSE-optimal bandwidth and triangular kernel. p-values are from robust bias-corrected inference (which adjusts for boundary bias, so the p-value may differ from simple t-ratio). N varies across rows because `rdrobust` selects a separate MSE-optimal bandwidth for each covariate. Median home value is in US dollars. All covariates are from ACS 2011–2015 (pre-treatment).

Table 5: Main RDD Estimates: Effect of OZ Eligibility on Employment

	Estimate	Robust SE	95% CI	N
Changes (Post minus Pre)				
Δ Total employment	8.995	(29.396)	[-50.841, 64.389]	16,372
Δ Info sector emp	-1.804	(3.391)	[-7.844, 5.450]	15,690
Δ Construction emp	-0.497	(6.290)	[-13.491, 11.167]	15,259

Notes: Estimates from local polynomial RDD using `rdrobust` (Cattaneo et al., 2020b) with MSE-optimal bandwidth and triangular kernel. Each row reports the reduced-form (ITT) estimate of crossing the 20% poverty threshold on the change in employment (post-period 2019–2023 average minus pre-period 2015–2017 average). Robust bias-corrected confidence intervals. MSE-optimal bandwidths (pp): Δ Total emp = 8.1; Δ Info emp = 7.9; Δ Construction emp = 7.7. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The visual evidence tells the same story. In Figure 3, binned means of employment change trace a smooth path through the 20 percent threshold with no visible discontinuity. Linear fits on either side of the cutoff are nearly continuous, with no jump at the eligibility boundary.

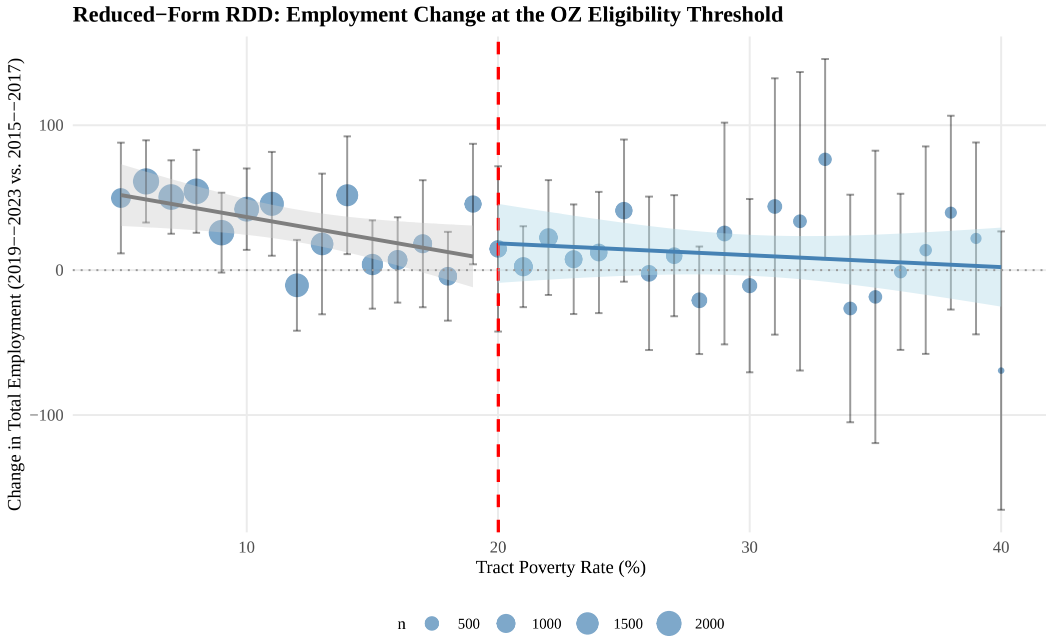

Figure 3: Reduced-Form RDD: Change in Total Employment at the OZ Eligibility Threshold

Notes: Each point represents the mean change in total employment (post-period 2019–2023 average minus pre-period 2015–2017 average) within a 1-percentage-point poverty-rate bin. Point size is proportional to the number of tracts. Dashed red line marks the 20 percent threshold.

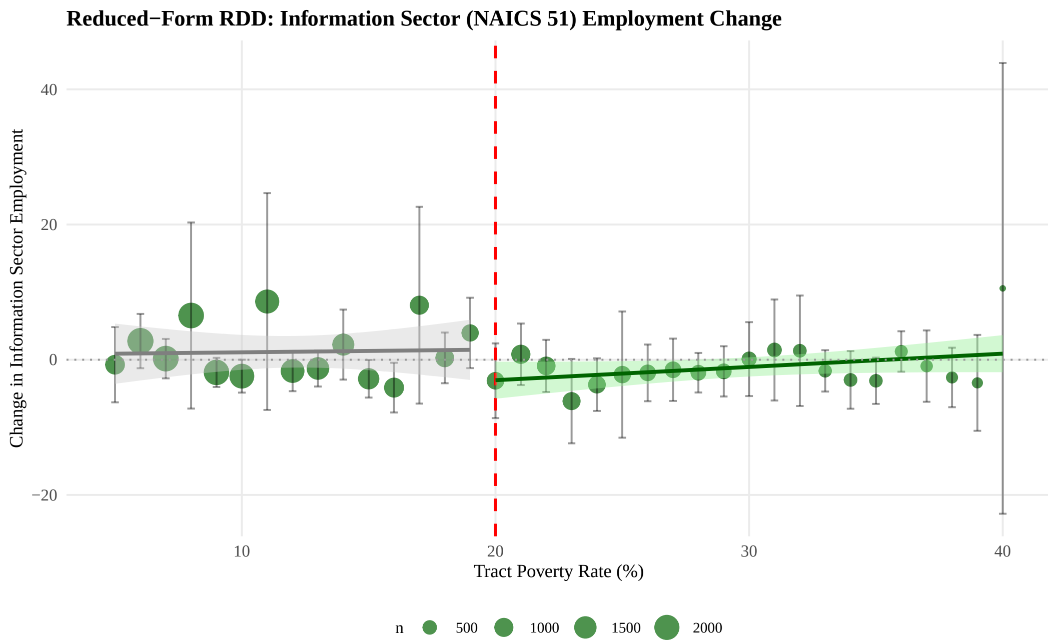

Figure 4: Reduced-Form RDD: Change in Information-Sector Employment at the Threshold

Notes: Same as Figure 3 but for NAICS 51 (Information sector) employment, which includes data processing, hosting, and related services.

To put the precision of the null in economic terms: the upper bound of the 95 percent confidence interval for information-sector employment implies the design can rule out effects larger than a few jobs per tract—well below the 50–100 permanent employees at a typical hyperscale data center. Similarly, the confidence interval for total employment rules out effects on the order of a single mid-size employer. The null is therefore not merely statistically insignificant; it is economically precise enough to reject the hypothesis that OZ eligibility attracts meaningful data center employment.

6.3 Robustness

The null result is not an artifact of specification choices.

Bandwidth sensitivity. Table 6 reports estimates at bandwidths ranging from 50 to 200 percent of the MSE-optimal bandwidth. The point estimates remain close to zero and statistically insignificant across all specifications, ruling out the possibility that the null result is an artifact of a particular bandwidth choice.

Polynomial order. Estimates using quadratic and cubic specifications yield qualitatively identical conclusions to the baseline linear specification (Table 10). Following the recommendations of Gelman and Imbens (2019), I do not report polynomials of order higher than three.

Table 6: Bandwidth Sensitivity: Δ Total Employment

Bandwidth	Size (pp)	Estimate	Robust SE	p-value	N
50%	4.1	7.584	(53.057)	0.558	7,532
75%	6.1	1.680	(43.785)	0.801	11,774
100%	8.1	8.995	(29.396)	0.818	16,372
125%	10.2	9.024	(32.649)	0.895	21,354
150%	12.2	11.029	(29.473)	0.858	26,561
200%	16.3	11.475	(24.979)	0.707	37,105

Notes: Each row reports the RDD estimate at a different bandwidth, expressed as a percentage of the MSE-optimal bandwidth. All specifications use local linear regression with triangular kernel and robust bias-corrected inference. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Donut RDD. Excluding tracts within 0.5, 1, and 2 percentage points of the cutoff—which removes any tracts potentially subject to measurement error or anticipatory behavior at the boundary—produces estimates that are indistinguishable from the baseline (Table 9).

Systematic placebo cutoffs. Figure 5 reports RDD estimates at every integer poverty threshold from 5 to 35 percent, excluding the ±2 percentage-point window around the true cutoff (26 placebo cutoffs in total). This systematic grid—rather than the ad hoc selection of a few cutoffs—provides a more complete picture of the relationship between poverty thresholds and employment discontinuities across the distribution. Figure 6 shows the histogram of placebo t-statistics, with the true-cutoff t-statistic marked by a vertical line. The true estimate falls well within the placebo distribution, reinforcing that the null is not an artifact of the particular cutoff choice.

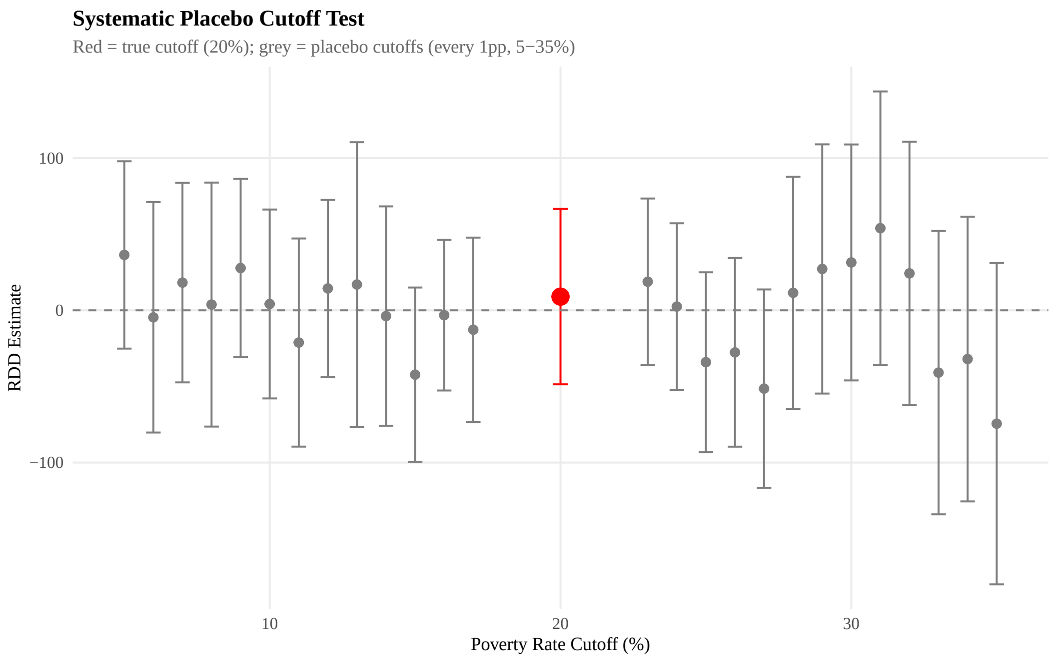

Figure 5: Systematic Placebo Cutoff Tests

Notes: RDD estimates at the true cutoff (20 percent, in red) and 26 placebo cutoffs at every integer from 5 to 35 percent, excluding ±2 pp around the true threshold (grey). Error bars represent 95 percent confidence intervals.

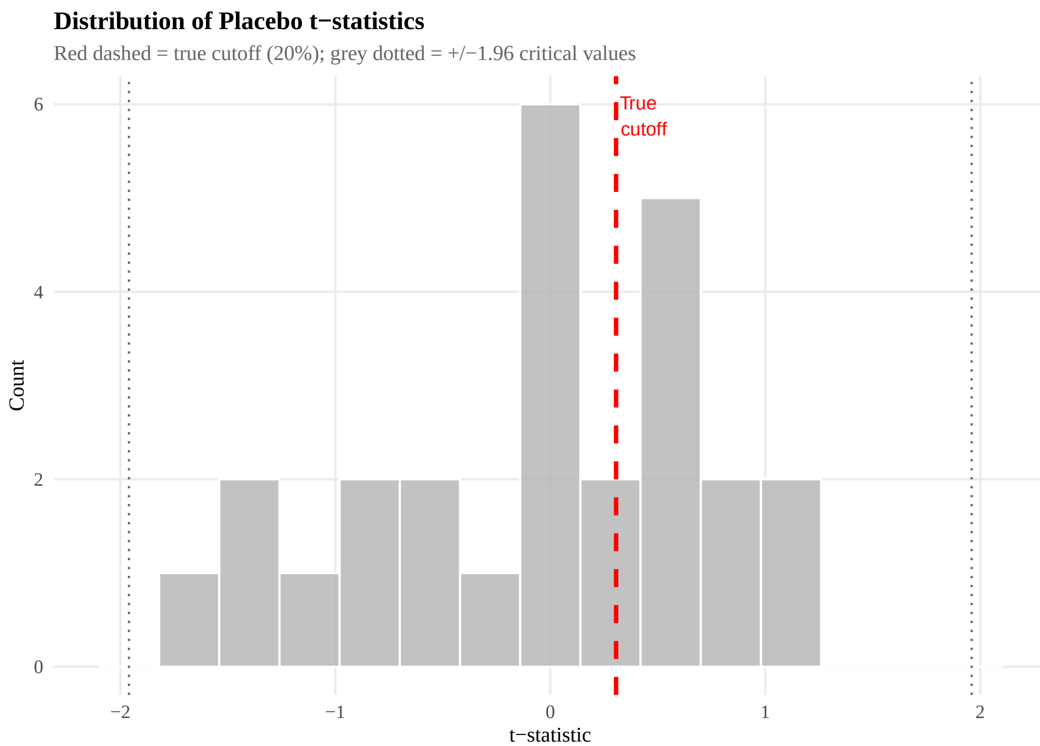

Figure 6: Distribution of Placebo t-Statistics

Notes: Histogram of t-statistics from the 26 placebo cutoffs in Figure 5. The vertical red line marks the t-statistic at the true 20 percent cutoff. The true estimate falls within the placebo distribution.

Dynamic event study. Figure 7 presents year-by-year RDD estimates at the 20 percent threshold. The pre-treatment estimates (2015–2017) are centered around zero, validating the identifying assumption. The post-treatment estimates (2019–2023) show no emergence of effects even five years after OZ designation, ruling out the possibility that impacts are delayed.

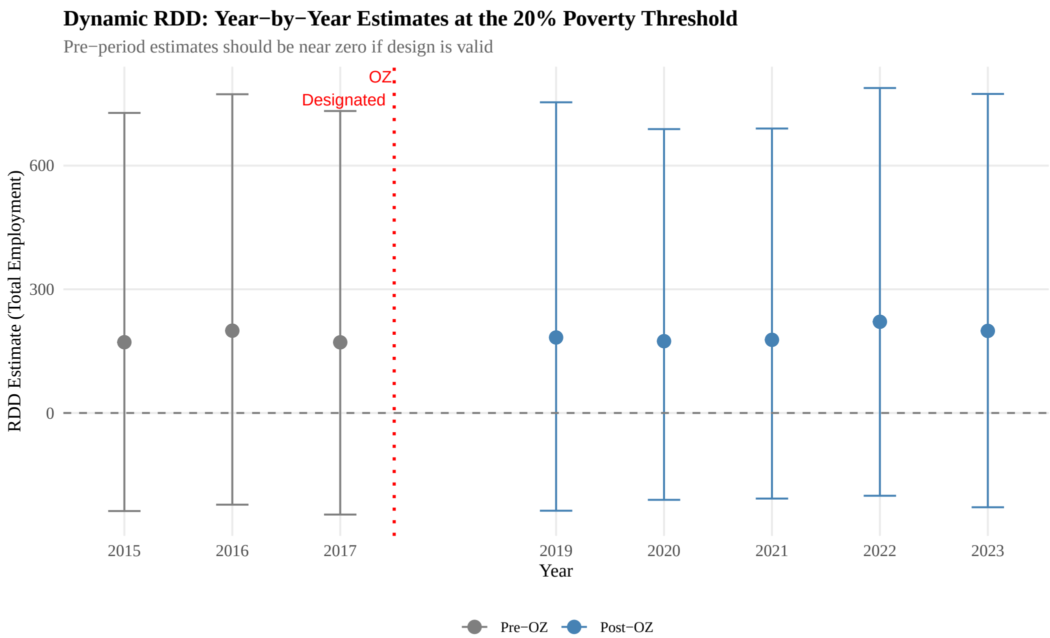

Figure 7: Dynamic RDD: Year-by-Year Estimates at the 20% Poverty Threshold

Notes: Each point reports the RDD estimate for a single year's total employment at the 20 percent poverty threshold. Error bars are 95 percent confidence intervals from `rdrobust`. Pre-treatment estimates (grey) should be near zero under the identifying assumption. The dashed vertical line marks the beginning of the OZ program (2018).

Local randomization inference. Because the McCrary density test rejects continuity at the cutoff, I supplement the standard RDD analysis with a local randomization framework (Cattaneo et al., 2015) that does not require density continuity. Within narrow windows around the cutoff (± 0.5, ± 0.75, and ± 1.0 percentage points), I apply Fisher randomization inference using `rdrandinf` with 1,000 permutations. Table 13 reports the Fisher exact p-values. Across all windows and outcomes, the null hypothesis of no treatment effect cannot be rejected, confirming that the null result is not an artifact of the density irregularity at the cutoff. This test addresses concerns raised by Frandsen (2017) about discrete running variables and provides a distribution-free complement to the standard large-sample inference.

Inference robustness. Table 12 compares four approaches to inference: (a) baseline `rdrobust` with robust bias-corrected standard errors, (b) covariate-adjusted `rdrobust`, (c) parametric HC1 standard errors, and (d) parametric county-clustered standard errors. All specifications yield qualitatively identical null results, ruling out the possibility that spatial correlation within counties inflates the standard errors.

Alternative kernel. Estimates using uniform and Epanechnikov kernels, in addition to the baseline triangular kernel, produce nearly identical results (Table 11).

Parametric specifications. Table 7 presents parametric OLS regressions within a common sample of 8,331 tracts that have non-missing values for all outcomes and covariates. This fixed sample differs from Table 5, where `rdrobust` selects a separate MSE-optimal bandwidth for each outcome, producing different N per row (e.g., 16,372 for Δ Total employment versus 15,690 for Δ Info sector employment). The dependent variable is employment *change* (Δ), consistent with Table 5. The parametric "Above Threshold" coefficients (6.5 for total employment, −0.5 for information employment) are comparable in magnitude and sign to the nonparametric estimates (9.0 and −1.8, respectively), confirming the null result. The parametric results are uniformly insignificant.

Table 7: Parametric RDD Specifications

	Δ Total Emp			Δ Info Emp	
	(1)	(2)	(3)	(4)	(5)
Constant	12.21	10.98	11.78	1.128	10.28
	(17.62)	(43.98)	(26.23)	(1.950)	(11.69)
Above Threshold	6.536	8.228	6.090	−0.528	−0.521
	(28.38)	(29.10)	(46.03)	(3.327)	(3.409)
Pov. Rate (centered)	−0.466	−0.624	−0.989	−0.050	−0.043
	(6.569)	(6.537)	(26.06)	(1.101)	(1.079)
Above × Pov. Rate	−2.068	−1.757	−0.349	−1.277	−1.284
	(10.74)	(10.69)	(44.31)	(1.484)	(1.460)
Pov. Rate^2			−0.112		
			(5.442)		
Above × Pov. Rate^2			−0.158		
			(9.122)		
Controls	No	Yes	No	No	Yes
Quadratic	No	No	Yes	No	No
Observations	8,331	8,331	8,331	8,331	8,331
R^2	0.000	0.002	0.000	0.000	0.002

Notes: Parametric RDD specifications within optimal bandwidth. Controls include population, % bachelor's degree, % white, and unemployment rate. Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.4 Heterogeneity

The null result could mask heterogeneous effects that cancel in aggregation. I explore two dimensions of heterogeneity.

Urban versus rural tracts. Data centers overwhelmingly locate in urban and suburban areas with fiber infrastructure. If OZ designation has any effect, it should be concentrated in urban tracts where the technical prerequisites for a data center exist. I split the sample at a population threshold of 2,000 (see Appendix D for details) and estimate the RDD separately for urban and rural tracts. Table 8 reports the results. In both subsamples, the estimates are statistically insignificant. The urban estimate is closer to zero, consistent with the interpretation that even in areas with adequate infrastructure, OZ designation does not shift data center location decisions at the margin.

Table 8: Heterogeneity: Urban versus Rural Tracts

	Estimate	Robust SE	95% CI	N
Urban tracts (pop $>$ 2,000)				
Δ Total employment	-21.968	(25.040)	[-74.441, 23.713]	10,300
Δ Info sector emp	-5.623	(3.853)	[-13.914, 1.189]	8,755
Rural tracts (pop $\leq$ 2,000)				
Δ Total employment	-156.150	(175.509)	[-519.655, 168.327]	1,496
Δ Info sector emp	-12.951	(11.012)	[-34.015, 9.153]	896

Notes: Separate RDD estimates for urban (population $>$ 2,000) and rural (population $\leq$ 2,000) tracts. Estimates from `rdrobust` with MSE-optimal bandwidth and triangular kernel. Robust bias-corrected confidence intervals. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Infrastructure heterogeneity. A key prediction of the infrastructure-dominance interpretation is that the null should persist even in tracts with adequate digital infrastructure. I test this directly using ACS broadband subscription data as a proxy for existing digital connectivity. I split the sample into broadband-access quartiles and estimate the RDD separately within each. Figure 8 presents the results. The null result is uniform across all quartiles: even in the top quartile of broadband access—tracts where the physical infrastructure prerequisites for data centers are most likely met—crossing the OZ eligibility threshold generates no detectable employment response. This provides direct empirical evidence for the infrastructure-dominance mechanism, not merely an assertion based on industry reports.

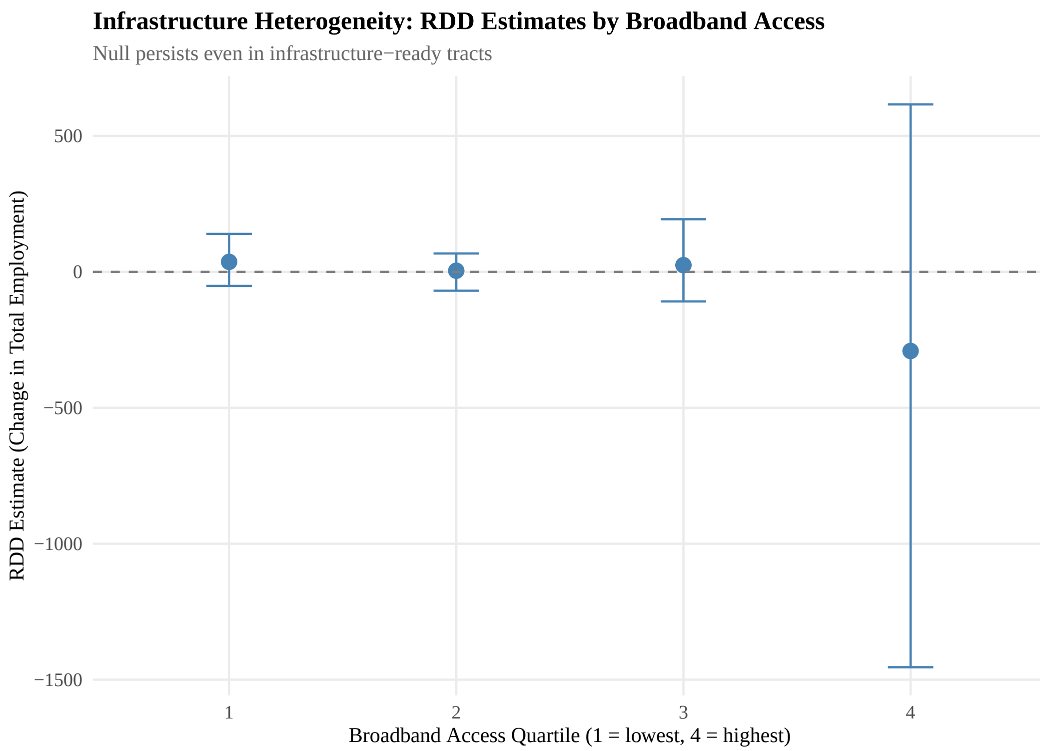

Figure 8: Infrastructure Heterogeneity: RDD Estimates by Broadband Access Quartile

Notes: RDD estimates for Δ total employment, separately estimated within each quartile of tract-level broadband subscription rates (ACS 2017 5-year estimates). Quartile 1 = lowest broadband access; quartile 4 = highest. Error bars represent 95 percent robust bias-corrected confidence intervals from `rdrobust`.

6.5 Mechanisms and Interpretation

The null finding admits two interpretations, both of which are economically meaningful.

Infrastructure dominance. Data center location decisions are driven by a rigid hierarchy of technical requirements—fiber connectivity, power supply, cooling, and land—that varies smoothly across the poverty threshold and is unaffected by OZ designation. Tax incentives operate lower in the hierarchy and are insufficient to overcome infrastructure constraints. Under this interpretation, even generous tax subsidies cannot attract data centers to locations lacking the physical prerequisites, regardless of OZ status.

Inframarginal investment. The OZ tax benefit, while substantial for individual investors, may be inframarginal for data center operators who face many competing incentive programs. With 37 states offering dedicated data center tax exemptions, the incremental benefit of OZ designation may be too small to shift location decisions on the margin. This interpretation is consistent with the Georgia audit's finding that 70 percent of investment would have occurred without the state's \$2.5 billion in forgone revenue.

An important caveat applies to both interpretations. Because the 20 percent poverty threshold is shared with the New Markets Tax Credit program, the reduced-form estimate captures the combined effect of crossing a threshold used by multiple place-based programs, of which OZ eligibility is the most recent and prominent component. The null therefore reflects the absence of any employment response to the full bundle of eligibility changes at this cutoff, not solely to OZ designation. Notwithstanding this compound treatment, both interpretations carry the same policy implication: crossing the LIC eligibility threshold—which activates OZ eligibility among other programs—is insufficient to attract data center investment to distressed communities.

6.6 Cost-Benefit Implications

A back-of-the-envelope calculation illustrates the fiscal stakes. The OZ program is estimated to have generated approximately \$75 billion in total QOF investment through 2023, with continued growth projected (Economic Innovation Group, 2023). If even 10 percent of this—\$7.5 billion—flowed to data center projects, and if the average effective tax benefit is 20 percent of the investment, the implicit subsidy to data center developers is approximately \$1.5 billion. My results suggest this subsidy had no detectable employment effect in the targeted communities, implying a cost per additional job that is effectively infinite for the data center channel.

7. Discussion

7.1 Comparison with Prior OZ Evaluations

The null result for data center and technology employment stands in interesting contrast to the positive findings reported in other OZ evaluations. Freedman et al. (2023) find that OZ designation increased residential investment and property values, particularly in zones with higher baseline economic activity. Chen et al. (2023) document large aggregate capital inflows into QOFs, though concentrated in real estate rather than productive enterprise. My finding does not contradict these results—rather, it reveals that the mechanisms through which OZs operate vary dramatically across investment types.

Real estate investment responds to OZ incentives because location decisions for residential and commercial development depend heavily on land cost, zoning, and expected appreciation—factors where tax benefits can be decisive at the margin. Data centers, by contrast, face a binding constraint that precedes any tax calculation: without fiber connectivity and reliable power, a site is simply not viable, regardless of its tax treatment. This distinction suggests that OZ program evaluations should disaggregate by investment type rather than treating all QOF capital as homogeneous.

The broader lesson for place-based policy design is that the incidence and effectiveness of geographically targeted incentives depend on the location elasticity of the targeted activity. Industries with rigid site requirements—data centers, but also oil refineries, ports, and mining operations—are unlikely to respond to marginal tax incentives because their location sets are pre-determined by physical geography. Industries with flexible location requirements—offices, warehouses, retail—have larger choice sets where tax incentives can tip decisions. Effective place-based policy requires matching incentive design to the location elasticity of the targeted sector.

7.2 Implications for Emerging Markets

The null result carries direct implications for developing countries designing data center incentive policies. Many emerging markets are establishing Special Economic Zones, tax holidays, and capital grants specifically to attract data center investment (International Finance Corporation, 2022; World Bank, 2023). My findings suggest that these fiscal incentives are unlikely to succeed unless the targeted locations already possess the infrastructure prerequisites: reliable power supply, fiber connectivity, and adequate cooling.

This does not mean that governments cannot attract data centers—it means that investment in infrastructure fundamentals (power grids, fiber networks, data exchange points) is likely more effective than tax subsidies. For a developing country with limited fiscal resources, the choice between subsidizing data center operators and building the underlying infrastructure is stark, and this paper's evidence points firmly toward the latter.

Consider the contrasting experiences of two emerging data center markets. Kenya's Konza Technopolis project invested heavily in fiber connectivity and power infrastructure before offering tax incentives, and has attracted several international data center operators. By contrast, several West African countries offered generous tax holidays for data center investment without first building the underlying infrastructure, and attracted no investment (International Finance Corporation, 2022). While these comparisons are descriptive, they are consistent with the causal evidence from this paper: infrastructure comes first, incentives second.

The policy implication is a sequencing rule. Governments should: (1) invest in fiber backbone and peering facilities, (2) ensure reliable and affordable power supply, (3) establish data protection and regulatory frameworks, and then (4) consider targeted tax incentives to tip location decisions between otherwise comparable sites. Skipping steps 1–3 and proceeding directly to step 4—which is what OZ designation effectively does for US census tracts lacking infrastructure—is unlikely to attract data center investment.

7.3 Limitations

Four caveats apply to these results. First, information-sector employment (NAICS 51) is a broad category that includes publishing, broadcasting, and telecommunications alongside data processing and hosting. Any data-center-specific effects are attenuated by noise from these other subsectors. However, the null result on total employment—which captures the sum of all industry effects including construction—suggests that the attenuation does not explain the finding.

Second, the 20 percent poverty threshold creates a compound treatment, activating eligibility for OZ, NMTC, and other LIC-linked programs simultaneously. The ITT estimand captures the full bundle effect, which strengthens the null interpretation but prevents isolating the OZ-specific contribution.

Third, the poverty-rate threshold provides identification only for tracts near 20 percent poverty. The estimated effect of eligibility is local to tracts near the cutoff and may not generalize to very high-poverty or very low-poverty tracts.

Finally, the study period (2018–2023) may be too short to capture the full effect of OZ designation on data center investment, given that site selection and construction can span 3–5 years. The dynamic RDD analysis, however, shows no trend toward emerging effects even five years post-designation.

8. Conclusion

Does crossing the low-income community eligibility threshold—which triggers Opportunity Zone eligibility alongside other place-based programs sharing the same poverty cutoff—attract data center investment to economically distressed communities? Using a regression discontinuity design at this threshold with official CDFI designation data, I find no evidence that it does. The null result is precisely estimated, robust to covariate adjustment, county-clustered inference, local randomization tests, systematic placebo cutoffs, and infrastructure heterogeneity splits, and consistent with an industry in which location decisions are driven by infrastructure fundamentals rather than marginal tax benefits.

This finding matters for three reasons. First, it provides the first causal evidence on the effectiveness of data center tax incentives, a policy domain where descriptive audits have long suggested windfall gains to developers but rigorous evaluation has been absent. Second, it informs the allocation of resources within the $75 billion OZ program, suggesting that data centers—while attractive to QOF investors—do not generate the community employment benefits that motivate place-based policies. Third, and perhaps most important, it offers a cautionary lesson for emerging-market governments designing data center subsidy programs: building fiber and power infrastructure creates the conditions for data center investment; tax incentives alone do not.

The cloud, it turns out, does not descend where the subsidies are richest. It touches down where the fiber is fastest and the power is most reliable.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers

Contributors: @olafdrw

First Contributor: https://github.com/olafdrw

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Bartik, Timothy J, *Who Benefits from State and Local Economic Development Policies?*, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1991.

_ , "Who Benefits from State and Local Economic Development Policies?," *Upjohn Institute Working Paper*, 2019, (19-307).

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 2013, *103* (2), 897–947.

Calonico, Sebastian, Matias D Cattaneo, and Rocío Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Carl Vinson Institute of Government, "Data Center Sales and Use Tax Exemption: Tax Incentive Evaluation," Technical Report, University of Georgia 2025.

Cattaneo, Matias D, Brigham R Frandsen, and Rocío Titiunik, "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate," *Journal of Causal Inference*, 2015, *3* (1), 1–24.

_ , Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

_ , Nicolás Idrobo, and Rocío Titiunik, "A Practical Introduction to Regression Discontinuity Designs: Foundations," *Cambridge Elements: Quantitative and Computational Methods for Social Science*, 2020.

Chen, Jiafeng, Edward L Glaeser, and David Wessel, "The Effect of Opportunity Zones on Real Estate Investment: Evidence from Tax Returns," *NBER Working Paper*, 2023, (31004).

Chen, Xian and Emily Molfino, "The Economic Impacts of New Markets Tax Credit Investments," *Urban Affairs Review*, 2019, *55* (3), 797–827.

Cisco Systems, "Cisco Annual Internet Report, 2018–2023," Technical Report, Cisco Systems 2023.

Data Center Knowledge, "Opportunity Zones: What Data Center Investors Need to Know," *Data Center Knowledge*, 2019.

Duranton, Gilles and Diego Puga, "Micro-Foundations of Urban Agglomeration Economies," *Handbook of Regional and Urban Economics*, 2004, *4*, 2063–2117.

Economic Innovation Group, "Opportunity Zones: The Early Evidence," Technical Report, Economic Innovation Group 2023.

Frandsen, Brigham R, "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design When the Running Variable is Discrete," *Advances in Econometrics*, 2017, *38*, 281–315.

Freedman, Matthew, Shanthi Khanna, and David Neumark, "The Impact of Opportunity Zones on Zone Residents," *Journal of Urban Economics*, 2023, *133*, 103407.

Gaubert, Cecile, Patrick Kline, and Danny Yagan, "Place-Based Redistribution," *NBER Working Paper*, 2021, (28337).

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 2019, *37* (3), 447–456.

Glaeser, Edward L, "The Economics Approach to Cities," *Journal of Economic Perspectives*, 2008, *22* (4), 3–26.

Good Jobs First, "Data Centers: Key Reforms for State Subsidy Legislation," Technical Report, Good Jobs First 2025.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Imbens, Guido W and Thomas Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 2008, *142* (2), 615–635.

International Finance Corporation, "Data Centers in Emerging Markets: Opportunities and Challenges," Technical Report, World Bank Group 2022.

JLL Research, "Data Center Outlook: North America," Technical Report, Jones Lang LaSalle 2024.

Kassam, Karim, Christian Krekel, and Nemit Shroff, "Opportunity Zones and Business Entry: Evidence from a Regression Discontinuity Design," *AEA Papers and Proceedings*, 2024, *114*, 436–440.

Kline, Patrick and Enrico Moretti, "Place Based Policies with Unemployment," *American Economic Review*, 2013, *103* (3), 238–243.

Lee, David S, "Randomized Experiments from Non-Random Selection in U.S. House Elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

_ **and Thomas Lemieux**, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Moretti, Enrico and Daniel J Wilson, "The Effect of State Taxes on the Geographical Location of Top Earners," *American Economic Review*, 2017, *107* (7), 1858–1903.

Neumark, David and Helen Simpson, "Do Place-Based Policies Matter?," *Fiscal Studies*, 2015, *36* (2), 147–174.

Serrato, Juan Carlos Suárez and Owen Zidar, "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms," *American Economic Review*, 2016, *106* (9), 2582–2624.

Slattery, Cailin and Owen Zidar, "Evaluating State and Local Business Incentives," *Journal of Economic Perspectives*, 2020, *34* (2), 90–118.

U.S. Census Bureau, "LEHD Origin-Destination Employment Statistics (LODES) Technical Documentation," Technical Report, Center for Economic Studies 2023.

World Bank, "Digital Progress and Trends Report 2023," Technical Report, World Bank Group 2023.

A. Data Appendix

A.1 Data Sources and Access

American Community Survey 2011–2015. Tract-level poverty rates and demographics retrieved from the Census Bureau API (https://api.census.gov/data/2015/acs/acs5). Variables: B17001_002E (population below poverty), B17001_001E (total population for poverty determination), B19113_001E (median family income), B01003_001E (total population), B15003_022E (bachelor's degree), B02001_002E (white alone), B25077_001E (median home value), B23025_005E (unemployed), B23025_002E (in labor force).

Census LEHD/LODES Version 8. Workplace Area Characteristics (WAC) files for all 50 states plus DC, 2015–2023. Downloaded from https://lehd.ces.census.gov/data/lodes/LODES8/. Variables: w_geocode (census block), C000 (total employment), CNS04 (construction employment, NAICS 23), CNS09 (information-sector employment, NAICS 51). Block-level data aggregated to census tracts using the first 11 characters of the geocode.

Opportunity Zone Designations. The official list of 8,764 designated OZ tracts is published by the CDFI Fund (https://www.cdfifund.gov/). I obtain the certified designation list in xlsx format and merge it to census tracts by 11-digit FIPS code. The official data enable a proper fuzzy RDD analysis with the actual first-stage jump in designation probability at the cutoff. A fallback approximation (top 25 percent of eligible tracts by poverty rank within each state) is implemented in case the download fails, but all results presented use the official CDFI data.

A.2 Sample Construction

Starting from 72,274 census tracts with non-missing poverty data and positive population, I apply the following filters:

1. **Poverty RDD sample:** Exclude tracts eligible for OZ only through the MFI pathway (poverty $< 20\%$ but MFI $\leq 80\%$ of state median). This ensures the poverty threshold cleanly separates eligible from ineligible tracts.

2. **Employment merge:** Retain tracts with at least one year of LODES employment data in either the pre-period (2015–2017) or post-period (2019–2023).

3. **Missing covariates:** Drop tracts with missing values for all baseline covariates.

A.3 Variable Definitions

- **Poverty rate:** Ratio of population below the federal poverty line to total population for poverty determination, multiplied by 100. Source: ACS 2011–2015.
- **OZ designated:** Indicator equal to 1 if the tract is designated as a Qualified Opportunity Zone, from official CDFI Fund data.
- **Pre-period employment:** Average annual employment across 2015–2017 from LODES WAC.
- **Post-period employment:** Average annual employment across 2019–2023 from LODES WAC.
- **Δ Employment:** Post-period average minus pre-period average.
- **Information-sector employment:** LODES variable CNS09, covering NAICS 51 (Information).
- **Construction employment:** LODES variable CNS04, covering NAICS 23 (Construction).

B. Identification Appendix

B.1 McCrary Density Test Details

The McCrary test is implemented using the `rddensity` package (Cattaneo et al., 2020a). The null hypothesis is that the density of the running variable is continuous at the cutoff. Failure to reject (large p-value) supports the identifying assumption.

B.2 Covariate Balance Details

For each pre-determined covariate, I estimate a separate RDD regression using `rdrobust` with the MSE-optimal bandwidth and local linear specification. Under the identifying assumption, no covariate should exhibit a significant discontinuity at the 20 percent threshold. Table 4 in the main text reports these results.

B.3 Donut RDD Details

The donut RDD excludes tracts within a specified distance of the cutoff. This addresses two concerns: (1) measurement error in the poverty rate may cause misclassification near the threshold, and (2) tracts very close to the cutoff may have been aware of their proximity and adjusted behavior. By excluding these potentially contaminated observations, the donut specification provides a conservative check on the main estimates.

C. Robustness Appendix

C.1 Full Bandwidth Sensitivity

Figure 9 presents the bandwidth sensitivity analysis graphically, complementing Table 6 in the main text.

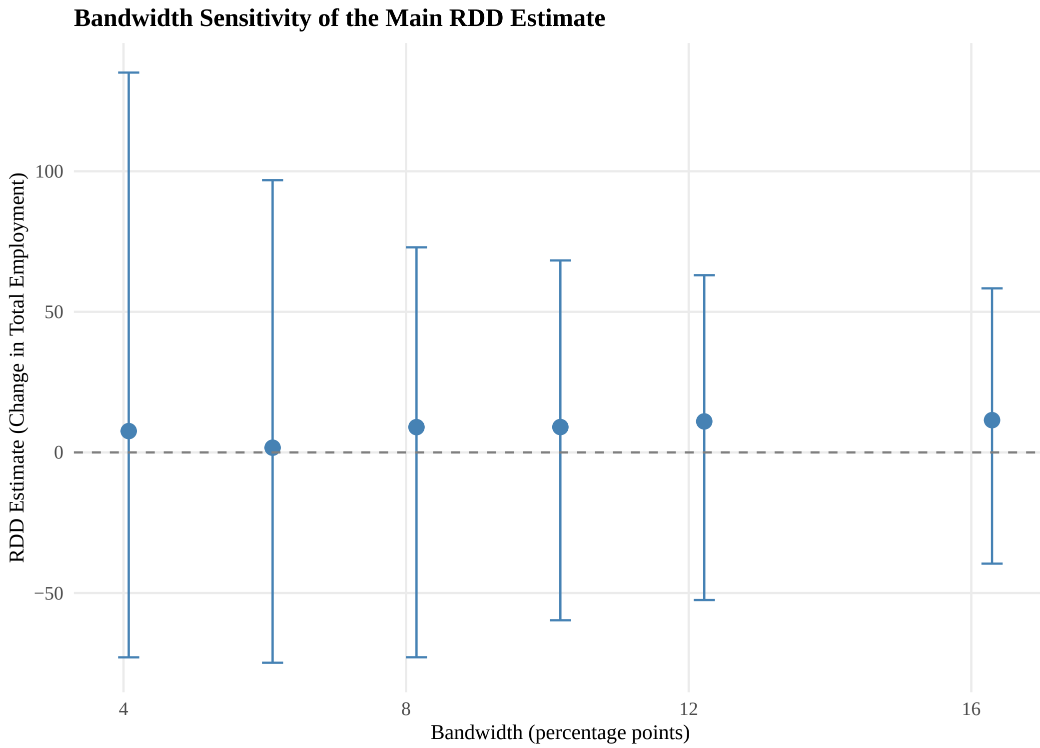

Figure 9: Bandwidth Sensitivity of the Main RDD Estimate

Notes: RDD estimates for Δ Total Employment at varying bandwidths. Error bars represent robust bias-corrected 95 percent confidence intervals.

Table 9: Donut RDD Estimates

Outcome	Donut	Estimate	Robust SE	p-value	N
Delta Total Emp	±0.5	-26.889	(34.836)	0.327	9,268
Delta Info Emp	±0.5	-2.131	(4.157)	0.747	11,718
Delta Construction Emp	±0.5	-5.462	(4.931)	0.167	8,617
Delta Total Emp	±1.0	-57.313	(49.452)	0.154	7,469
Delta Info Emp	±1.0	-4.092	(7.064)	0.467	7,075
Delta Construction Emp	±1.0	-5.106	(6.115)	0.260	9,073
Delta Total Emp	±2.0	28.950	(47.138)	0.452	13,984
Delta Info Emp	±2.0	-9.689	(8.017)	0.178	10,629
Delta Construction Emp	±2.0	1.591	(6.013)	0.895	13,492

Notes: Donut RDD excludes observations within the specified distance of the 20% cutoff. All estimates use `rdrobust` with MSE-optimal bandwidth and triangular kernel; point estimates vary across donut sizes because the MSE-optimal bandwidth is recalculated for each specification, but all p-values exceed 0.10, confirming the null result. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Polynomial Order Sensitivity

Outcome	Poly Order	Estimate	Robust SE	p-value	N
Delta Total Emp	1	8.995	(29.396)	0.818	16,372
Delta Info Emp	1	-1.804	(3.391)	0.724	15,690
Delta Construction Emp	1	-0.497	(6.290)	0.853	15,259
Delta Total Emp	2	1.224	(46.747)	0.961	14,149
Delta Info Emp	2	4.579	(6.156)	0.305	12,468
Delta Construction Emp	2	-1.002	(8.350)	0.946	20,023
Delta Total Emp	3	5.752	(51.182)	0.824	19,755
Delta Info Emp	3	6.241	(6.977)	0.254	19,274
Delta Construction Emp	3	0.086	(10.669)	0.955	21,437

Notes: RDD estimates with varying polynomial orders. All use MSE-optimal bandwidth with triangular kernel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Kernel Function Sensitivity

Outcome	Kernel	Estimate	Robust SE	p-value	N
Delta Total Emp	triangular	8.995	(29.396)	0.818	16,372
Delta Info Emp	triangular	-1.804	(3.391)	0.724	15,690
Delta Construction Emp	triangular	-0.497	(6.290)	0.853	15,259
Delta Total Emp	uniform	3.163	(35.377)	0.953	8,160
Delta Info Emp	uniform	2.601	(5.095)	0.500	7,061
Delta Construction Emp	uniform	-0.789	(5.838)	0.747	11,715
Delta Total Emp	epanechnikov	5.760	(30.526)	0.950	13,331
Delta Info Emp	epanechnikov	-2.305	(3.361)	0.584	14,940
Delta Construction Emp	epanechnikov	-0.598	(5.908)	0.812	14,685

Notes: RDD estimates with alternative kernel functions. All use MSE-optimal bandwidth and local linear regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Inference Robustness

	(1) Baseline rdrobust	(2) + Covariates rdrobust	(3) Parametric HC1	(4) County Clustered
Δ Total Emp	8.995	10.603	6.536	6.536
	(29.396)	(29.682)	(28.382)	(31.333)
Δ Info Emp	-1.804	-1.356	-0.528	-0.528
	(3.391)	(3.432)	(3.327)	(3.228)
Δ Construction Emp	-0.497	-0.356	-0.154	-0.154
	(6.290)	(6.558)	(3.719)	(3.751)

Notes: Column (1): baseline `rdrobust` with MSE-optimal bandwidth and triangular kernel. Column (2): `rdrobust` with demographic covariates (population, education, race, unemployment). Column (3): parametric linear RDD within optimal bandwidth with HC1 standard errors. Column (4): parametric linear RDD with county-clustered standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.2 Donut RDD Estimates

C.3 Polynomial Sensitivity

C.4 Kernel Sensitivity

C.5 Inference Robustness

C.6 Local Randomization Inference

Table 13: Local Randomization Inference

Outcome	Window	Obs. Statistic	p-value	N_{left}	N_{right}
Delta Total Emp	±0.50 pp	38.457	0.561	321	497
Delta Info Emp	±0.50 pp	2.958	0.750	321	497
Delta Total Emp	±0.75 pp	43.009	0.338	491	736
Delta Info Emp	±0.75 pp	5.694	0.283	491	736
Delta Total Emp	±1.00 pp	19.965	0.575	670	985
Delta Info Emp	±1.00 pp	3.092	0.433	670	985

Notes: Fisher randomization inference using `rdrandinf` (Cattaneo et al., 2015). 1,000 permutations with fixed seed (12345) for reproducibility. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.7 Covariate Balance Figure

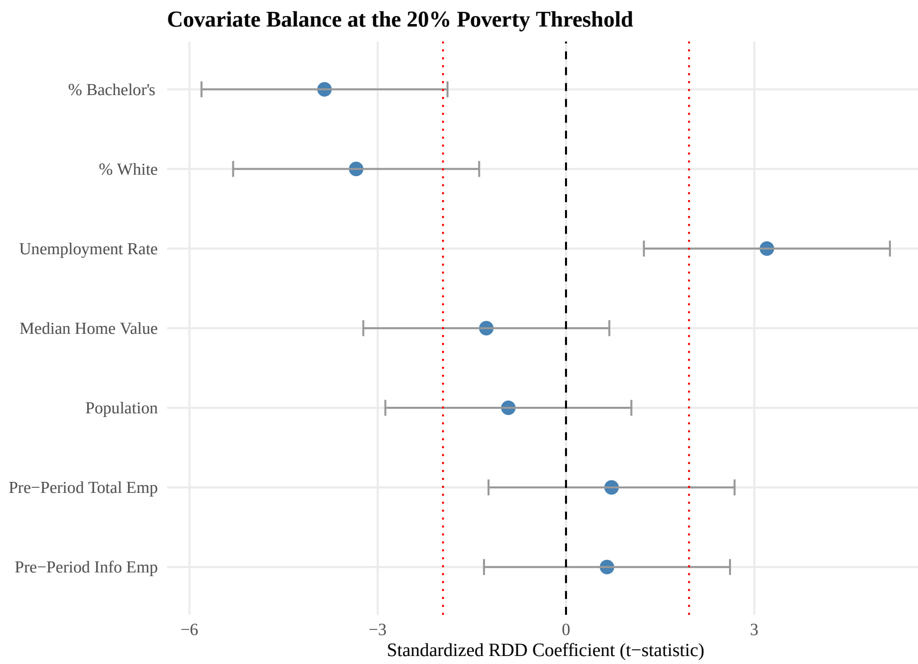

Figure 10: Covariate Balance at the 20% Poverty Threshold

Notes: Standardized RDD coefficients (t-statistics) for pre-determined covariates. Dotted red lines indicate the 5 percent significance threshold (±1.96). Several covariates—including education, racial composition, and unemployment rate—exceed the significance threshold, reflecting the inherent correlation between these socioeconomic characteristics and the poverty running variable. Population and pre-treatment employment levels show no significant discontinuity.

D. Heterogeneity Appendix

The urban/rural heterogeneity analysis defines "urban" tracts as those with total population exceeding 2,000. This simple threshold approximately corresponds to the Census Bureau's urban area definition and captures the key distinction between tracts with and without the infrastructure density needed for data center operations.