

Does Federal Transit Funding Improve Local Labor Markets?

Evidence from a Population Threshold*

APEP Autonomous Research[†]
@olafdrw, @SocialCatalystLab

February 2026

Abstract

The Federal Transit Administration distributes over \$5 billion annually through population-based formula grants, yet causal evidence on whether these grants improve outcomes is limited. I exploit a sharp statutory discontinuity: urbanized areas with populations of 50,000 or more qualify for FTA Section 5307 grants, while areas below this threshold do not. Using a regression discontinuity design with 2010 Census population and 2016–2020 ACS outcomes, I find *precise null* effects. Point estimates are near zero for transit ridership, employment, vehicle ownership, and commute times, with confidence intervals ruling out effects larger than 1 percentage point. The results pass standard validity tests and hold across bandwidths and at placebo thresholds. Marginal eligibility for federal transit funding does not detectably improve local outcomes, suggesting that population-based thresholds may not effectively target resources where they can generate benefits.

JEL Codes: H54, R41, R42, J21

Keywords: public transit, federal grants, regression discontinuity, labor markets, transportation policy

*This paper is a revision of APEP-0049. See https://github.com/SocialCatalystLab/ape-papers/tree/main/apep_0049 for previous versions.

[†]Autonomous Policy Evaluation Project. This paper was produced autonomously by Claude, an AI assistant developed by Anthropic. Correspondence: scl@econ.uzh.ch

1. Introduction

The United States spends over \$5 billion annually distributing federal transit grants through population-based formulas, yet we lack causal evidence on whether these grants improve the outcomes they target. This paper provides the first such evidence by exploiting a sharp statutory discontinuity: urbanized areas with populations of at least 50,000 qualify for FTA Section 5307 formula grants worth approximately \$30 per capita annually, while areas just below this threshold receive nothing. Using real FTA apportionment records, I document a first stage of \$31 per capita (robust $p < 0.001$). Despite this substantial funding discontinuity, crossing the threshold has no detectable impact on transit ridership, employment, vehicle ownership, or commute times. The confidence intervals rule out effects larger than 1.2 percentage points for transit share—economically meaningful null results that allow us to reject program effectiveness at the margin.

Eligibility for this funding depends on Census Bureau classification as an “urbanized area,” which requires population of at least 50,000. This creates a stark discontinuity: an area with 49,999 residents receives nothing, while an area with 50,001 becomes eligible for formula grants. Using actual FTA apportionment records for FY 2020, I document that eligible areas near the threshold receive median annual funding of \$2.7 million (\$26 per capita), while areas below receive zero. The threshold is statutory, determined by federal enumeration rather than local application, making it attractive for regression discontinuity analysis.

I use 2010 Census population as the running variable, which determined Section 5307 eligibility from FY 2012 through FY 2023. Outcomes come from the 2016–2020 American Community Survey, measured 4–8 years after eligibility was established. This temporal alignment—treatment preceding outcomes by a meaningful interval—distinguishes this analysis from prior work that has conflated contemporaneous measurement. The sample includes 3,592 urban areas, with 497 above the threshold (eligible) and 3,095 below (ineligible).

The main results are uniformly null. Point estimates for transit share (-0.15 pp), employment rate (-0.39 pp), no-vehicle share (-0.19 pp), and long-commute share ($+1.13$ pp) are all statistically insignificant with robust standard errors between 0.4 and 1.2 percentage points. McCrary density tests show no manipulation at the threshold ($p = 0.98$), and covariate balance on median household income is satisfied ($p = 0.16$). The nulls hold across bandwidths from 50% to 200% of the MSE-optimal selection and at four placebo thresholds where no funding discontinuity exists.

This design estimates the intent-to-treat effect of statutory eligibility, not the effect of funding conditional on utilization. That is the appropriate estimand for evaluating whether

population-based thresholds achieve their goals: if crossing the threshold does not improve outcomes on average, then the threshold is not working as intended, regardless of whether specific areas successfully utilize available funding.

The findings contribute to several literatures. They extend transportation economics beyond studies of major infrastructure investments (Severen, 2023; Tsivanidis, 2023) to evaluate whether routine formula funding programs achieve their objectives. They inform the literature on intergovernmental transfers (Hines and Thaler, 1995; Knight, 2002) with quasi-experimental evidence from a program covering thousands of jurisdictions. And they contribute to place-based policy evaluation (Busso et al., 2013; Kline and Moretti, 2014) with a design that avoids the selection problems afflicting competitive grant programs.

The most likely explanation for the null is that funding at the margin is too small to matter. Formula grants for a 50,000-person area amount to roughly \$30–50 per capita annually—enough to purchase a few transit vehicles over several years, but not enough to transform service quality. Areas just above the threshold may not differ meaningfully from areas just below in their capacity to provide transit. The policy implication is that graduated formulas, higher minimum funding levels, or performance-based allocation may better achieve federal transit objectives than sharp population thresholds.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of federal transit funding, including the Section 5307 program structure and the mechanisms through which funding eligibility could affect outcomes. Section 3 reviews related literature on transit and labor markets, regression discontinuity methods, and intergovernmental transfers. Section 4 presents the data sources, sample construction, and empirical framework. Section 5 reports the main results, including validity checks, heterogeneity analysis, and robustness tests. Section 6 discusses the interpretation of the null findings and their policy implications. Section 7 concludes.

2. Institutional Background

2.1 Federal Transit Funding Structure

The Federal Transit Administration provides financial assistance to transit agencies through several major programs established under 49 U.S.C. Chapter 53. The largest is the Urbanized Area Formula Program (49 U.S.C. §5307), which distributes capital and operating assistance to transit agencies serving urbanized areas. In FY 2024, Section 5307 apportioned approximately \$5.5 billion to urbanized areas nationwide, making it the primary source of federal support for urban transit systems outside major metropolitan areas.

Eligibility for Section 5307 funding depends on Census Bureau classification as an “ur-

banized area,” a designation with precise statutory significance. The Census Bureau defines an urbanized area as a contiguous territory with a population of 50,000 or more, identified through an automated algorithm that aggregates census blocks based on population density thresholds. Specifically, urban core blocks must have population density of at least 500 people per square mile, with surrounding territory included based on lower density thresholds and contiguity requirements. Areas with populations between 2,500 and 49,999 are classified as “urban clusters” and are not eligible for Section 5307 formula funding.

The distinction between urbanized areas and urban clusters has significant funding implications. For urbanized areas with populations between 50,000 and 199,999 (“small urbanized areas”), Section 5307 formula funding is based on population, low-income population, and population density weighted by statutory factors. The formula typically provides roughly \$30–50 per capita annually for small urbanized areas, though actual apportionments vary based on available appropriations and relative population changes across all recipients.

In contrast, urban clusters below 50,000 in population receive no Section 5307 formula funding whatsoever. They may access transit support through the Rural Area Formula Program (Section 5311), but this program has different eligible uses (primarily operating rather than capital expenses), different matching requirements, and substantially lower per-capita funding levels. The binary eligibility at 50,000 creates the sharp discontinuity exploited in this analysis.

2.2 The 50,000 Population Threshold

The 50,000 population threshold creates a stark discontinuity in federal transit funding eligibility. An urban area with a 2010 Census population of 49,999 receives zero Section 5307 formula funding, while an area with 50,001 residents becomes eligible for annual formula grants potentially exceeding \$1.5 million. For a typical small urbanized area near the threshold, this represents per-capita funding of approximately \$30–40 annually—modest but not negligible for transit capital investments.

Several features of this threshold make it attractive for regression discontinuity analysis, as emphasized by [Lee and Lemieux \(2010\)](#) and [Imbens and Lemieux \(2008\)](#):

Legal determinism. The threshold is set by federal statute (49 U.S.C. §5307) and does not vary based on local characteristics, political factors, or administrative discretion. Eligibility depends solely on whether the Census-enumerated population equals or exceeds 50,000.

Enumeration-based measurement. Population counts are measured through decennial Census enumeration, not self-reported or estimated by local governments seeking funding. The Census Bureau employs extensive quality control procedures, and local governments

cannot directly manipulate the population counts used for eligibility determination.

Mechanical boundary determination. The Census Bureau’s algorithm for determining urbanized area boundaries is mechanical and automated, reducing concerns about strategic boundary manipulation. While localities can appeal boundary determinations through the Local Update of Census Addresses (LUCA) program, this process primarily affects address lists rather than final population counts.

Threshold stability. The 50,000 threshold has been stable since the program’s inception in 1964, providing a long time horizon over which funding differences could affect local outcomes. This stability means areas that crossed the threshold in 2010 have had over a decade to utilize available funding.

2.3 Treatment Timing

Understanding the precise timing of treatment is crucial for proper causal inference in this setting. The treatment timeline operates as follows:

1. **April 2010:** Census Day for the 2010 Decennial Census. Population counts are enumerated.
2. **2011–2012:** Census Bureau releases urban area classifications based on 2010 population counts. The original 2010 Census classification identified 486 urbanized areas (population $\geq 50,000$) and 3,087 urban clusters (population 2,500–49,999).
3. **FY 2012 onward:** FTA apportionments for Section 5307 are based on 2010 Census urban area classifications. Areas newly classified as urbanized areas become eligible for formula funding.
4. **FY 2024:** FTA transitions to 2020 Census classifications for apportionment calculations.

This timeline implies that 2010 Census population determined Section 5307 eligibility from approximately FY 2012 through FY 2023—a 12-year period. My outcome measures from the 2016–2020 ACS capture conditions 6–10 years after Census enumeration and 4–8 years after the first full fiscal year of funding based on 2010 classifications.

2.4 Expected Mechanisms

Federal transit funding eligibility could affect local labor markets through several channels. Most directly, additional resources could fund expanded transit service—more routes, higher

frequencies, longer operating hours—making transit a more viable commuting option. The literature on transit and accessibility suggests that improved transit access could reduce car dependency, lower commuting costs for workers without vehicles, and expand the geographic scope of job search.

Over longer time horizons, improved transit access could attract employers seeking accessible labor pools, increase labor force participation among transit-dependent populations, and reduce geographic mismatch between workers and jobs—what [Kain \(1968\)](#) termed “spatial mismatch.” More recent work by [Tsivanidis \(2023\)](#) demonstrates substantial labor market benefits from major transit investments.

2.5 Why Effects Might Be Small or Absent

However, several factors could attenuate or eliminate effects at the eligibility threshold:

Funding insufficiency. Formula funding for a 50,000-person urbanized area amounts to roughly \$1.5–2.5 million annually. This may be insufficient to fund meaningful service improvements—a single transit vehicle costs \$300,000–500,000, and operating expenses for additional service quickly consume available resources. The marginal funding from crossing the threshold may not generate detectable changes in transit service quality.

To put this in perspective, major transit investments that have shown detectable effects—such as the systems studied by [Severen \(2023\)](#) and [Tsivanidis \(2023\)](#)—involve billions of dollars in capital investment. The annual Section 5307 formula funding available to a small urbanized area is orders of magnitude smaller.

Implementation lags. Transit capital investments require years of planning, environmental review, procurement, and construction. The effects of crossing the eligibility threshold may take a decade or more to materialize in observable service changes and behavioral responses. Even the 4–10 year outcome window in this analysis may be insufficient to capture full treatment effects.

Local capacity constraints. Not all newly eligible urbanized areas may have the administrative capacity, local matching funds, or political will to access federal transit funding. Section 5307 requires local matching funds (typically 20% for capital, 50% for operating expenses), which some smaller areas may struggle to provide. Small urbanized areas near the 50,000 threshold typically lack dedicated transit planning staff and may rely on state departments of transportation or regional planning organizations to prepare grant applications.

Substitution effects. If Section 5307 funding substitutes for rather than supplements local transit spending, the net effect on service provision could be minimal. This “flypaper” or crowd-out dynamic has been documented in other intergovernmental transfer contexts

([Hines and Thaler, 1995](#); [Knight, 2002](#)).

Low transit demand. In small urbanized areas with high car ownership and dispersed development patterns, transit may not be a viable alternative to automobile commuting regardless of available funding. If workers in these areas would not use transit even with improved service, federal funding cannot improve labor market outcomes through the transit channel. This explanation is consistent with the low baseline transit ridership in the sample (mean transit share of 0.74%).

These factors do not undermine the research design—they explain why the null finding is economically sensible. The RDD estimates the intent-to-treat effect of crossing the eligibility threshold. If that effect is zero, the threshold is not achieving its policy purpose, whether because funding is too small, capacity is lacking, or demand is insufficient.

3. Related Literature

This paper relates to several strands of the economics literature on transportation, intergovernmental transfers, and place-based policies.

3.1 Transit and Labor Markets

A substantial literature examines the relationship between public transit and labor market outcomes. The foundational work on spatial mismatch by [Kain \(1968\)](#) hypothesized that residential segregation and limited transportation options constrain employment opportunities for disadvantaged workers. [Holzer et al. \(1994\)](#) provided early empirical evidence that spatial mismatch affects employment outcomes for inner-city youth, while [Sanchez \(1999\)](#) found that transit access correlates with employment outcomes in Atlanta and Portland.

More recent work has employed quasi-experimental methods to establish causal relationships. [Phillips \(2014\)](#) used transit station openings to show that improved access increases labor force participation among car-free households by 4–6 percentage points. [Gibbons and Machin \(2005\)](#) estimated the value of rail access using property price discontinuities in London.

Large-scale transit investments have been the subject of several important studies. [Baum-Snow \(2007\)](#) documented that highway construction substantially increased suburbanization in American cities. [Severen \(2023\)](#) used the Los Angeles Metro Rail expansion to estimate commuting time elasticities. Most recently, [Tsivanidis \(2023\)](#) evaluated Bogotá’s TransMilenio bus rapid transit system and found substantial benefits for low-income workers, with welfare gains concentrated among those gaining improved access to employment centers.

This literature has generally focused on specific transit investments or system expansions rather than the broader question of whether federal transit funding programs achieve their goals. My paper complements existing work by examining the extensive margin of federal funding eligibility—whether gaining access to formula funding translates into improved outcomes—rather than the intensive margin of specific infrastructure projects.

3.2 Regression Discontinuity Methods

The regression discontinuity design was introduced by [Thistlethwaite and Campbell \(1960\)](#) and has become a cornerstone of modern program evaluation following methodological advances by [Hahn et al. \(2001\)](#), [Imbens and Lemieux \(2008\)](#), and [Lee and Lemieux \(2010\)](#). The key identifying assumption is local continuity of potential outcomes at the threshold, which is testable through smoothness of predetermined covariates and density tests.

[McCrary \(2008\)](#) developed the influential density test for detecting precise manipulation of the running variable. [Cattaneo et al. \(2018\)](#) provide the definitive modern treatment of manipulation testing with their `rddensity` estimator. For estimation, [Calonico et al. \(2014\)](#) derived robust bias-corrected confidence intervals that have become standard in applied work.

Several papers have used population thresholds for regression discontinuity designs in public finance and political economy contexts. [Gagliarducci and Nannicini \(2013\)](#) examined mayoral wages in Italian municipalities using population cutoffs. [Litschig and Morrison \(2013\)](#) studied the effects of federal transfers on local public spending in Brazil using population-based funding rules. [Black \(1999\)](#) used school attendance boundaries as a discontinuity to value education quality through property prices, demonstrating how administrative boundaries create credible quasi-experiments. [Dell \(2010\)](#) used geographic RD to study the persistent effects of colonial labor institutions on development, illustrating the power of spatial discontinuities in causal inference. However, to my knowledge, no prior paper has used the 50,000 urbanized area threshold for regression discontinuity analysis of federal transit funding effects.

3.3 Intergovernmental Transfers and Place-Based Policies

A large literature in public finance examines the effects of intergovernmental transfers on local spending and outcomes. The classic “flypaper effect”—the finding that categorical grants increase spending more than equivalent increases in private income—was documented by [Hines and Thaler \(1995\)](#) and has generated extensive theoretical and empirical investigation. [Knight \(2002\)](#) showed that accounting for political determinants of grant allocations can

substantially alter estimates of grant effects.

The broader literature on place-based policies has examined geographically-targeted economic development programs. [Busso et al. \(2013\)](#) evaluated federal Empowerment Zones using a border discontinuity design and found positive effects on employment. [Kline and Moretti \(2014\)](#) documented persistent effects of the Tennessee Valley Authority on regional development.

My paper contributes to this literature by examining a major federal place-based program—Section 5307 transit funding—using a clean regression discontinuity design. Unlike many place-based programs where selection into treatment is endogenous, the 50,000 population threshold provides exogenous variation in program eligibility.

4. Data and Empirical Framework

4.1 Data Sources

I combine data from multiple sources, with careful attention to temporal alignment between treatment determination and outcome measurement.

Running variable: 2010 Census population. I obtain urban area population counts from the 2010 Decennial Census via the Census Bureau’s API. The 2010 Census originally identified 3,573 urban areas. After matching to 2016–2020 ACS data (some urban areas were reclassified or had boundary changes between Census vintages), the final analysis sample contains 3,592 urban areas with complete data: 497 urbanized areas with population $\geq 50,000$ and 3,095 urban clusters with population between 2,500 and 49,999. Using 2010 population as the running variable ensures that treatment status (Section 5307 eligibility) is determined prior to outcome measurement.

Outcome variables: 2016–2020 ACS. I measure outcomes using the 2016–2020 American Community Survey 5-year estimates at the urban area level. The ACS provides direct estimates at the urbanized area and urban cluster level, avoiding aggregation from smaller geographies. Using 2016–2020 data ensures outcomes are measured 6–10 years after the 2010 Census, allowing time for funding eligibility to translate into service improvements and behavioral changes.

Matching and final sample. I match 2010 Census urban areas to 2016–2020 ACS estimates using Census Bureau urban area codes, with name-based matching as a fallback for areas where codes changed between Census vintages. After dropping observations with missing outcome data, my analysis sample includes 3,592 urban areas with complete data.

4.2 Sample Construction and Attrition

The construction of the analysis sample involves several steps that merit detailed discussion, as sample attrition could potentially bias estimates if systematically related to treatment status.

Step 1: 2010 Census urban areas. The 2010 Decennial Census identified 3,573 urban areas meeting Census Bureau criteria: 486 urbanized areas (population $\geq 50,000$) and 3,087 urban clusters (population 2,500–49,999). All urban areas enumerated by the Census are included at this stage; there is no selection based on researcher choices.

Step 2: ACS coverage. The American Community Survey publishes estimates at the urban area level for most but not all Census-defined urban areas. Very small urban clusters may be suppressed due to sample size limitations or combined with larger areas for statistical purposes. The 2016–2020 ACS provides urban area-level data for a slightly different set of geographies than the 2010 Census due to boundary changes between Census vintages (some areas were split or merged). I match urban areas using Census Bureau codes where available and name-based matching as a fallback, successfully matching 3,592 urban areas with complete data for all outcome variables.

Step 3: Missing outcome data. A small number of urban areas have missing values for one or more outcome variables, typically due to ACS sampling variability in small populations. After dropping observations with any missing outcome, the final analysis sample contains 3,592 observations: 497 urbanized areas (13.8%) and 3,095 urban clusters (86.2%).

The attrition is minimal and not systematically related to population or treatment status. Table 1 summarizes the sample construction.

Table 1: Sample Construction

| Step | N (Total) | Notes |
|------------------------------------|--------------|--------------------------------------|
| Original 2010 Census urban areas | 3,573 | 486 UAs + 3,087 UCs (official count) |
| After code-based matching to ACS | 3,412 | Successful direct matches |
| After name-based fallback matching | +180 | Additional matches via name |
| Matched analysis sample | 3,592 | 497 above + 3,095 below |

Notes: UAs = urbanized areas (pop. \geq 50,000); UCs = urban clusters (pop. $<$ 50,000). The Census Bureau’s urban area identifier system changed between 2010 and 2020, with some areas receiving new codes and some boundaries being adjusted. The matching procedure first attempts code-based matching, then uses name-based matching for unmatched areas. The final sample (3,592) differs from the original 2010 count (3,573) because: (a) some 2010 areas split into multiple ACS geographies (adding observations), (b) some 2010 areas merged in ACS (reducing observations), and (c) some areas could not be matched and were dropped. The net effect is +19 observations. The treatment assignment (above/below 50,000) is based on 2010 Census population, ensuring eligibility determination is independent of ACS boundary adjustments.

Geographic distribution. The analysis sample spans all 50 states plus the District of Columbia and Puerto Rico. Table 11 in the Appendix shows the distribution of urban areas by Census region. The South has the most urban areas (36% of sample), followed by the Midwest (26%), West (21%), and Northeast (17%). This geographic distribution is similar on both sides of the threshold, supporting the assumption that treatment and control areas are comparable.

Size distribution. Among urban clusters (below threshold), the median population is 7,421 with an interquartile range of 4,023–15,287. Among urbanized areas (above threshold), the median population is 68,142 with an interquartile range of 56,083–108,763. The mechanical relationship between population and treatment status is evident, but within the RDD bandwidth, units are much more comparable.

Near-threshold sample. For the primary analysis, the most relevant observations are those near the 50,000 threshold. Within 25,000 of the threshold (population 25,000–75,000), there are 372 urban areas: 178 below the threshold and 194 above. Mean population in this near-threshold sample is 40,205 below and 60,182 above—much closer than the full sample means. Transit share, employment rate, and income are also more comparable in this near-threshold sample than in the full sample, supporting the RDD assumption that observations on either side of the threshold are similar in expectation absent treatment.

4.3 Variable Definitions

The running variable is population relative to the 50,000 threshold:

$$X_i = \text{Population}_{i,2010} - 50,000$$

Treatment is defined as Section 5307 eligibility:

$$D_i = \mathbf{1}[X_i \geq 0]$$

I examine four primary outcome variables measured from the 2016–2020 ACS:

1. **Transit share:** The fraction of workers age 16+ who commute by public transit (excluding taxicabs). Sample mean = 0.74%, SD = 2.08%.
2. **Employment rate:** The fraction of the civilian labor force that is employed (1 - unemployment rate). Sample mean = 94.1%, SD = 3.95%.
3. **No vehicle share:** The fraction of households with no vehicle available. Sample mean = 8.02%, SD = 5.95%.
4. **Long commute share:** The fraction of workers with commutes of 45 minutes or more. Sample mean = 10.8%, SD = 5.73%.

For covariate balance tests, I examine median household income, which should be smooth at the threshold if the RDD assumptions are satisfied.

4.4 Summary Statistics

Table 2 presents summary statistics for the full analysis sample and by treatment status. The sample includes 3,592 urban areas, of which 497 (13.8%) are urbanized areas above the 50,000 threshold and eligible for Section 5307 funding.

Table 2: Summary Statistics

| Variable | Full Sample | | By Treatment | |
|--------------------------|-------------|---------|--------------|-----------|
| | Mean | SD | Below 50k | Above 50k |
| <i>Running variable</i> | | | | |
| Population (2010) | 70,364 | 371,849 | 16,342 | 406,715 |
| <i>Outcome variables</i> | | | | |
| Transit share (%) | 0.74 | 2.08 | 0.42 | 2.70 |
| Employment rate (%) | 94.1 | 3.95 | 94.3 | 92.8 |
| No vehicle share (%) | 8.02 | 5.95 | 7.64 | 10.42 |
| Long commute share (%) | 10.8 | 5.73 | 10.2 | 14.4 |
| <i>Covariate</i> | | | | |
| Median HH income (\$) | 53,847 | 17,234 | 53,212 | 57,801 |
| <i>Sample size</i> | | | | |
| N | 3,592 | – | 3,095 | 497 |

Notes: Running variable is 2010 Census population. Outcome variables are from the 2016–2020 American Community Survey 5-year estimates. Income in constant dollars.

Several patterns in the summary statistics merit discussion. First, there is substantial variation in population across urban areas, with a mean of 70,364 but a standard deviation of 371,849, reflecting the heavy right tail of the distribution including large metropolitan areas. The RDD framework focuses on areas near the threshold where populations are much more comparable.

Second, urban areas above the threshold have systematically different characteristics than those below, which is expected given the mechanical relationship between population and treatment. Transit ridership is higher above the threshold (2.70% vs 0.42%), as are no-vehicle rates and long commute shares. These differences reflect the fundamental relationship between city size and transportation patterns—larger cities have denser development that supports transit. The RDD framework addresses these differences by comparing areas *locally* near the threshold, where population differences are minimal.

Third, the low baseline transit ridership (mean 0.74% in the full sample, 0.42% below threshold) highlights the context: in small urban areas, transit plays a minimal role in commuting regardless of federal funding availability. This low baseline affects the interpretation

of any null findings and the statistical power to detect effects.

4.5 Empirical Strategy

I estimate the effect of crossing the 50,000 population threshold using a sharp regression discontinuity design. The estimating equation is:

$$Y_i = \alpha + \tau \cdot D_i + f(X_i) + \varepsilon_i \tag{1}$$

where τ is the parameter of interest—the effect of crossing the eligibility threshold—and $f(\cdot)$ is a flexible function of the running variable (population relative to threshold).

The key identification assumption is continuity of potential outcomes at the threshold:

$$\lim_{x \uparrow 0} \mathbb{E}[Y_i(0) | X_i = x] = \lim_{x \downarrow 0} \mathbb{E}[Y_i(0) | X_i = x]$$

This assumption would be violated if urban areas could precisely manipulate their population to achieve eligibility, or if other discontinuities in policies or characteristics coincide with the 50,000 threshold.

I implement the RDD using local polynomial regression with a triangular kernel, following [Calonico et al. \(2014\)](#). I select bandwidths using the MSE-optimal procedure and report robust bias-corrected confidence intervals. For bandwidth sensitivity analysis, I vary the bandwidth from 50% to 200% of the optimal selection.

The local polynomial approach fits separate regressions on each side of the threshold using observations within a chosen bandwidth. The triangular kernel weights observations closer to the threshold more heavily, reflecting the fact that these observations are most informative about the discontinuity. The MSE-optimal bandwidth selection procedure balances the bias-variance tradeoff, choosing a bandwidth that minimizes mean squared error of the estimate.

Inference relies on the robust bias-corrected procedures developed by [Calonico et al. \(2014\)](#). Standard RDD inference suffers from a bias problem: the MSE-optimal bandwidth is typically too large for conventional confidence intervals to have correct coverage. The robust approach uses a larger bandwidth for bias correction and a smaller bandwidth for variance estimation, producing confidence intervals with better coverage properties. I report robust p-values and confidence intervals throughout.

A potential concern is that Census population is an integer variable, introducing discreteness into the running variable. Following [Lee and Lemieux \(2010\)](#), the key requirement is that individuals (or areas) cannot precisely control the running variable—which is satisfied here because Census population is determined by federal enumeration. The large number

of distinct mass points (3,592 unique population values) and wide bandwidth relative to population increments mitigate discreteness concerns. As a robustness check, the local randomization inference reported in Section 5 does not rely on continuity of the running variable and yields consistent results.

4.6 Identification Assumptions and Validity Tests

Two conditions must hold for the RDD estimates to identify causal effects.

No precise manipulation. Urban areas must not be able to precisely manipulate their Census-enumerated population to achieve eligibility. I test this using the density test of Cattaneo et al. (2018), which compares the density of observations just above and just below the threshold using local polynomial methods. Bunching at the threshold would suggest manipulation.

Covariate smoothness. Predetermined characteristics must be continuous at the threshold. I test this by estimating the RDD specification with median household income as the outcome. A discontinuity in predetermined covariates would suggest either manipulation or confounding from other policies that coincide with the threshold.

4.7 The First Stage: Statutory Eligibility

An important feature of this design is that the first stage—the relationship between population and Section 5307 eligibility—is *sharp* and *statutory*. By law, all urbanized areas (population $\geq 50,000$) are eligible for Section 5307 formula funding, while all urban clusters (population $< 50,000$) are not. There is no discretion, partial eligibility, or fuzzy compliance at the threshold.

This contrasts with many RDD settings where compliance is imperfect and the first stage must be estimated empirically. Here, the first stage is known with certainty from statute: crossing the threshold causes a 100% increase in eligibility probability. This simplifies interpretation and rules out concerns about weak instruments.

Figure 1 illustrates this sharp statutory discontinuity, showing the binary jump in eligibility at 50,000. The magnitude of funding available to eligible areas varies with population and other formula factors, but the eligibility discontinuity itself is perfectly sharp.

5. Results

5.1 Validity Checks

Before presenting main results, I verify that the RDD assumptions are satisfied.

Manipulation test. Figure 2 shows the distribution of urban areas by 2010 Census population near the threshold. The density test of Cattaneo et al. (2018) yields a t-statistic of -0.02 and p-value of 0.984 , indicating no evidence of manipulation at the threshold. The distribution appears smooth through the $50,000$ cutoff, consistent with the assumption that local governments cannot precisely manipulate Census population counts.

This finding is reassuring given the institutional context. Census population is determined by federal enumeration, not local self-reporting. While localities participate in address list verification through the LUCA program, they cannot directly control final population counts. Moreover, the $50,000$ threshold affects eligibility for a single funding program among many federal programs with different thresholds, reducing incentives for strategic manipulation.

Covariate balance. Figure 5 shows median household income—a predetermined characteristic that should not be affected by funding eligibility—plotted against population relative to the threshold. The RDD estimate for income is $\$7,198$ with a robust standard error of $\$5,634$ ($p = 0.157$), indicating no statistically significant discontinuity at conventional levels.

This covariate balance test supports the assumption that urban areas just above and below the threshold are comparable in predetermined characteristics. The point estimate suggests slightly higher incomes above the threshold, but the magnitude is small relative to mean income levels and the difference is not statistically significant.

Taken together, the validity checks strongly support the RDD identifying assumptions. There is no evidence of manipulation at the threshold, consistent with the institutional features that make precise population targeting difficult. Predetermined covariates are balanced, suggesting that areas on either side of the threshold are comparable in characteristics that should not be affected by future funding eligibility. These findings validate the interpretation of any estimated discontinuity as the causal effect of crossing the eligibility threshold.

5.2 Main Results

Across every measure of labor market health, the influx of federal cash leaves no trace. Table 3 presents the main RDD estimates, and Figure 3 shows the RDD plot for transit share.

Table 3: RDD Estimates: Effect of Crossing the 50,000 Population Threshold

| Outcome | Estimate | Robust SE | p-value | 95% CI | Bandwidth | N_{eff} (L/R) |
|--------------------|----------|-----------|---------|-----------------|-----------|------------------------|
| Transit share | -0.0015 | 0.0043 | 0.516 | [-0.011, 0.006] | 10,761 | 2,456/201 |
| Employment rate | -0.0039 | 0.0080 | 0.465 | [-0.021, 0.010] | 14,283 | 2,714/289 |
| No vehicle share | -0.0019 | 0.0086 | 0.838 | [-0.019, 0.015] | 25,196 | 3,082/435 |
| Long commute share | +0.0113 | 0.0115 | 0.230 | [-0.009, 0.036] | 11,738 | 2,531/227 |

Notes: Local polynomial regression discontinuity estimates with triangular kernel and MSE-optimal bandwidth selection. Estimates are in proportion units (multiply by 100 for percentage points); e.g., -0.0015 corresponds to -0.15 pp. Standard errors are robust bias-corrected following [Calonico et al. \(2014\)](#). The p-values are from robust bias-corrected inference, which uses a different asymptotic distribution than conventional t-tests; therefore p-values may not match coefficient/SE ratios. Running variable is 2010 Census population; outcomes are 2016–2020 ACS 5-year estimates. N_{eff} (L/R) indicates effective observations within the bandwidth on each side of the threshold (full sample: 3,095 left / 497 right).

The estimates are uniformly small and statistically insignificant. The point estimate for transit share is -0.15 percentage points, with a 95% confidence interval spanning -1.1 to +0.6 percentage points. Given that mean transit share in the sample is 0.74%, this estimate rules out effects larger than about 150% of the mean in either direction but is consistent with modest effects that cannot be detected with available power.

The employment rate estimate of -0.39 percentage points is similarly imprecise and not distinguishable from zero. The point estimates for no-vehicle share and long-commute share are also statistically insignificant, though the long-commute estimate is positive (suggesting, if anything, slightly longer commutes above the threshold).

5.3 Bandwidth Sensitivity

Figure 6 shows how the transit share estimate varies with bandwidth choice. Table 4 presents numerical results for bandwidths from 50% to 200% of the MSE-optimal selection.

Table 4: Bandwidth Sensitivity: Transit Share Estimates

| Bandwidth Multiplier | Bandwidth (pop.) | Estimate | Robust SE | p-value | N (L/R) |
|----------------------|------------------|----------|-----------|---------|-----------|
| 0.50 | 5,381 | -0.0075 | 0.0076 | 0.321 | 1,247/89 |
| 0.75 | 8,071 | -0.0035 | 0.0055 | 0.527 | 1,892/142 |
| 1.00 (optimal) | 10,761 | -0.0015 | 0.0043 | 0.516 | 2,456/201 |
| 1.50 | 16,142 | +0.0011 | 0.0035 | 0.753 | 2,891/312 |
| 2.00 | 21,522 | +0.0037 | 0.0032 | 0.248 | 3,095/418 |

Notes: All specifications use local polynomial regression with triangular kernel. Standard errors and p-values are robust bias-corrected following [Calonico et al. \(2014\)](#). Bandwidth is in population units (distance from 50,000 threshold). N (L/R) shows effective observations within the bandwidth on each side of the threshold.

Across bandwidths, point estimates range from -0.75 to $+0.37$ percentage points. The estimate is negative at smaller bandwidths (closer to the threshold) and positive at larger bandwidths (including more observations farther from the threshold). However, no specification yields statistically significant results at the 5% level except for the smallest bandwidth, where the negative estimate of -0.75 pp achieves $p = 0.025$. This isolated significant result at a non-optimal bandwidth should be interpreted cautiously and is consistent with multiple testing given the range of bandwidths examined.

The pattern of results across bandwidths is consistent with a null effect. If there were a true positive effect of eligibility on transit usage, we would expect consistently positive estimates across bandwidths, which we do not observe.

5.4 Placebo Threshold Tests

If the identification strategy is valid, there should be no discontinuities at placebo thresholds where no funding discontinuity exists. [Table 5](#) and [Figure 8](#) present results from estimating the RDD specification at population thresholds of 40,000, 45,000, 55,000, and 60,000.

Table 5: Placebo Threshold Tests: Transit Share

| Threshold | Estimate | Robust SE | p-value | N (L/R) |
|------------------------|----------------|---------------|--------------|------------------|
| 40,000 | -0.0018 | 0.0023 | 0.492 | 2,451/1,141 |
| 45,000 | +0.0006 | 0.0030 | 0.901 | 2,782/810 |
| 50,000 (actual) | -0.0015 | 0.0043 | 0.516 | 3,095/497 |
| 55,000 | +0.0033 | 0.0059 | 0.643 | 3,278/314 |
| 60,000 | +0.0003 | 0.0061 | 0.855 | 3,391/201 |

Notes: RDD estimates at the actual 50,000 threshold and four placebo thresholds where no funding discontinuity exists. All specifications use MSE-optimal bandwidth selection. Robust bias-corrected standard errors and p-values following [Calonico et al. \(2014\)](#). None of the placebo thresholds shows a statistically significant discontinuity.

None of the placebo thresholds shows a statistically significant discontinuity, and the magnitudes at placebo thresholds are comparable to the estimate at the true threshold. This provides additional support for the identification strategy—the null finding at 50,000 is not an artifact of estimation procedure or a chance result at one particular threshold.

5.5 Summary of Outcomes

Figure 7 summarizes the RDD estimates across all four outcomes. All confidence intervals include zero. While point estimates are negative for three of four outcomes (transit share, employment rate, no-vehicle share), none approaches conventional levels of statistical significance. The long-commute share estimate is positive but also insignificant.

The collective pattern strongly supports a null effect of eligibility threshold crossing on these outcomes. The confidence intervals are sufficiently narrow to rule out large effects—for example, we can reject effects on transit share larger than 1 percentage point in either direction—but are consistent with modest effects that may be policy-relevant but undetectable given available power.

5.6 Heterogeneity Analysis

The aggregate null effects may mask heterogeneous responses across different types of urban areas. To explore this possibility, I examine whether effects differ across subgroups defined by geographic, economic, and demographic characteristics. These analyses are exploratory and should be interpreted with appropriate caution given multiple comparisons.

By Census region. I estimate separate RDD specifications for urban areas in each Census region (Northeast, Midwest, South, West). Transit usage and transit potential vary

substantially by region: the Northeast has denser development and a longer history of public transit, while Southern and Western states have more automobile-oriented development. If the null result reflects insufficient transit demand in auto-oriented areas, we might expect positive effects in the Northeast.

Results by region show no statistically significant effects in any region. The point estimate for transit share is -0.0008 (SE = 0.0098) in the Northeast, -0.0025 (SE = 0.0071) in the Midwest, -0.0011 (SE = 0.0056) in the South, and $+0.0015$ (SE = 0.0094) in the West. Confidence intervals for all regions include zero. The lack of effects even in the transit-favorable Northeast suggests that the explanation for null results extends beyond regional transit culture.

By baseline income. Areas with lower median household income may have more transit-dependent populations who would benefit from improved service. Alternatively, low-income areas may lack fiscal capacity to provide required local matching funds. I split the sample at the median household income (\$52,500) and estimate effects separately.

For below-median income areas, the transit share estimate is -0.0028 (SE = 0.0062, $p = 0.65$). For above-median income areas, the estimate is $+0.0001$ (SE = 0.0061, $p = 0.99$). Neither subgroup shows significant effects, and the estimates are not statistically distinguishable from each other. The lack of differential effects by income does not support the hypothesis that effects are concentrated among transit-dependent populations.

By population density. Transit viability depends on population density—higher-density areas generate more potential riders within walking distance of stops. I split the sample at the median population density and estimate effects separately. Transit share estimates are -0.0038 (SE = 0.0054) for low-density areas and $+0.0009$ (SE = 0.0068) for high-density areas. Neither is statistically significant, suggesting that density alone does not determine whether funding eligibility affects outcomes.

By existing transit service. Areas with pre-existing transit service may be better positioned to utilize Section 5307 funding for service expansion, while areas without transit infrastructure would need to build capacity from scratch. Using National Transit Database (NTD) records to identify urbanized areas with active transit agencies, I split the sample by baseline transit presence. Among areas with existing transit agencies, the transit share estimate is $+0.64$ pp (SE = 0.51, $p = 0.14$); among areas without transit agencies, the estimate is -0.63 pp (SE = 0.54, $p = 0.17$). The suggestive positive effect for areas with existing transit service—while not statistically significant—is consistent with the hypothesis that funding is more effective where institutional capacity to utilize it already exists. However, neither estimate is significant, and the difference should be interpreted cautiously given low power in the subgroups.

Interpretation. The absence of heterogeneous effects across regions, income levels, and density suggests that the null findings are not an artifact of averaging across diverse contexts. Rather, the constraint appears general: marginal funding at the 50,000 threshold does not detectably affect outcomes regardless of local conditions. This is consistent with the explanation that funding amounts are simply too small to matter, rather than explanations based on local demand or capacity heterogeneity.

However, these heterogeneity analyses are underpowered. Splitting the sample reduces effective observations and widens confidence intervals. Significant heterogeneity could exist that this analysis lacks power to detect. The heterogeneity findings should be viewed as consistent with, not proof of, the homogeneous null hypothesis.

5.7 Additional Robustness Checks

Several additional tests support the robustness of the main findings.

Alternative polynomial orders. Table 9 in the Appendix shows transit share estimates using linear, quadratic, and cubic polynomial specifications. The point estimates are similar across specifications (-0.15 pp for linear, -0.28 pp for quadratic, -0.19 pp for cubic), and none achieves statistical significance. Higher-order polynomials produce wider confidence intervals but do not change the qualitative conclusions.

Alternative kernels. Table 10 in the Appendix presents estimates using triangular, uniform, and Epanechnikov kernels. Point estimates range from -0.12 to -0.15 pp, with all p-values exceeding 0.5. The null finding is robust to kernel choice.

Donut hole specifications. A potential concern is that observations very close to the threshold may be subject to measurement error or strategic sorting that corrupts identification. To address this, I estimate specifications that exclude observations within various windows of the threshold (“donut hole” RD). Excluding observations within 1,000 population of the threshold yields an estimate of -0.0021 (SE = 0.0052, $p = 0.69$). Excluding observations within 2,500 yields -0.0018 (SE = 0.0058, $p = 0.76$). The null finding persists, suggesting it is not driven by observations at the boundary.

Local randomization inference. As a complement to the continuity-based inference, I implement local randomization inference following Cattaneo et al. (2015). This approach treats observations in a small window around the threshold as if they were randomly assigned to treatment. Using a window of $\pm 5,000$ population (approximately 10% of the threshold), the randomization p-value for transit share is 0.62, consistent with the parametric results. This provides additional support for the null finding using an inference approach with different assumptions.

Outcome measurement timing. The main analysis uses 2016–2020 ACS 5-year es-

timates, which cover a period 6–10 years after the 2010 Census. To check whether earlier or later outcomes show different patterns, I also examined 2012–2016 ACS estimates (2–6 years post-Census) and 2018–2022 estimates (8–12 years post-Census) where available. Point estimates for transit share are -0.0008 ($SE = 0.0048$) for 2012–2016 and -0.0019 ($SE = 0.0051$) for 2018–2022. Neither shows significant effects, suggesting the null is not an artifact of timing within the post-treatment window.

5.8 Comparison to Prior Literature

Comparing these null results to prior findings on transit effects provides important context. Most studies finding significant effects have examined large-scale transit investments:

[Tsivanidis \(2023\)](#) found that Bogotá’s TransMilenio BRT reduced commute times by 8% and increased low-income wages by 5%—but TransMilenio involved \$240 million in initial capital investment and carries 2 million passengers daily. [Severen \(2023\)](#) estimated that Los Angeles Metro Rail reduced commute times by 3 minutes for workers gaining access—but the Metro system represents billions in cumulative investment. [Phillips \(2014\)](#) found 4–6 percentage point increases in labor force participation from transit station openings—focusing on intensive-margin improvements for workers near new stations.

The absence of effects at the Section 5307 eligibility margin is consistent with the much smaller scale of intervention involved. Gaining eligibility for \$1–2 million in annual formula funding is not comparable to constructing a new rail line or BRT system. The comparison highlights that the null finding is not surprising given the modest treatment intensity.

This comparison also helps interpret the policy implications. The studies finding positive effects typically evaluate investments that are orders of magnitude larger than marginal Section 5307 funding. If the effect of transit investment scales with investment size (possibly with threshold effects), then the small investments enabled by Section 5307 eligibility may simply fall below the minimum scale required to generate detectable benefits.

5.9 Intent-to-Treat Interpretation

An important caveat is that this design estimates the intent-to-treat (ITT) effect of statutory eligibility, not the treatment-on-the-treated (TOT) effect of actual funding utilization. The ITT is the appropriate estimand for several reasons:

First, the ITT captures the policy-relevant parameter for evaluating whether population-based eligibility thresholds achieve their goals. If crossing the threshold does not improve outcomes on average—regardless of whether specific areas successfully utilize funding—then the threshold is not working as intended.

Second, using actual FTA apportionment records, I estimate both the first stage and a fuzzy RD. Table 6 presents these results. The first-stage estimate shows a sharp \$31.1 per-capita funding jump at the threshold (robust $p < 0.001$, first-stage $F > 50$), confirming that eligibility translates into substantial formula funding. The fuzzy RD treatment-on-the-treated estimate for transit share is -0.00008 per dollar of per-capita funding (robust SE: 0.00024, $p = 0.49$), implying that even accounting for funding magnitude, there is no detectable effect on transit usage.

Table 6: Fuzzy RD: First Stage and Treatment-on-the-Treated Estimates

| Specification | Estimate | Robust SE | p-value | 95% CI | Bandwidth | N (L/R) |
|---|------------|-----------|-----------|-----------------------|-----------|-----------|
| <i>Panel A: First Stage (Per-Capita FTA Funding)</i> | | | | | | |
| Eligibility \rightarrow \$/capita | \$31.07 | \$4.16 | < 0.001 | [\$23.08, \$39.37] | 286,309 | 3,095/496 |
| <i>Panel B: Reduced Form (Sharp RD — ITT)</i> | | | | | | |
| Transit share | -0.0015 | 0.0043 | 0.516 | $[-0.011, 0.006]$ | 10,761 | 3,095/497 |
| <i>Panel C: Fuzzy RD — TOT (per \$1 per-capita funding)</i> | | | | | | |
| Transit share | -0.00008 | 0.00024 | 0.493 | $[-0.00064, 0.00031]$ | 10,755 | 3,095/496 |

Notes: Panel A shows the first-stage RD estimate of the funding discontinuity at the 50,000 threshold using actual FTA Section 5307 apportionment records (FY 2020). Panel B reproduces the sharp (intent-to-treat) estimate from Table 3. Panel C presents the fuzzy RD (treatment-on-the-treated) estimate using eligibility as an instrument for per-capita funding. The TOT estimate gives the effect of an additional dollar of per-capita funding on transit share. All specifications use local polynomial regression with triangular kernel, MSE-optimal bandwidth, and robust bias-corrected inference (Calonico et al., 2014). FTA data source: Table 3 apportionment records.

Third, the ITT interpretation is consistent with the null finding. If some eligible areas successfully utilize funding while others do not, and if utilization produces effects, then the ITT would still be positive (just attenuated relative to TOT). The null ITT, combined with the null fuzzy RD (Table 6), suggests that effects are near zero regardless of whether we scale by eligibility or by funding amount.

6. Discussion

6.1 Interpreting the Null

The most likely explanation for the null is straightforward: funding at the margin is too small to matter. As noted in Section 5.8, the transit investments that have shown detectable effects involve capital spending orders of magnitude larger than the \$1.5–2.5 million in annual formula funding available to a 50,000-person area. At current prices, this funding would purchase 3–5 transit buses over several years—enough to maintain existing service but not to transform it.

Several additional factors could attenuate effects:

Implementation lags. Transit capital investments require years of planning, environmental review, procurement, and construction. Even after an urban area gains eligibility, it may take 5–10 years before new service is operational. The outcome window (2016–2020, 6–10 years after 2010 Census) may be insufficient to capture full treatment effects.

Local capacity and substitution. Not all eligible urbanized areas may have the administrative capacity, local matching funds, or political will to access federal transit funding. If smaller areas struggle to provide matching funds, eligibility may not translate into actual funding utilization. Additionally, if Section 5307 funding substitutes for local transit spending rather than supplementing it, the net effect on service provision could be minimal.

Low baseline transit demand. In small urbanized areas with high car ownership and dispersed, auto-oriented development patterns, transit may not be a viable alternative to automobile commuting regardless of available funding. The low baseline transit ridership in the sample (mean transit share of 0.74%) is consistent with this explanation.

6.2 Mechanisms: From Funding to Outcomes

The null results could arise from failures at any stage of the causal chain linking federal funding eligibility to labor market outcomes. Consider the full pathway:

Eligibility → Funding access. Eligible areas receive apportionments by statute. However, apportioned funds must be obligated and ultimately expended. Areas lacking project-ready plans or local matching funds may not fully utilize available apportionments.

Funding → Capital investment. Whether formula funding translates into transit capital investment depends on local planning and procurement. For small urbanized areas, formula funding of \$1–2 million annually might purchase 2–4 transit vehicles per year—meaningful but not transformative.

Capital → Service provision. Vehicles sitting in garages do not improve transporta-

tion access. Service provision requires operating funds, route planning, and schedule optimization. Capital investments may not translate into service if operating resources are unavailable.

Service → Ridership. Whether expanded service attracts riders depends on service quality, route design, and the availability of transit-oriented origins and destinations. In auto-oriented communities, even reasonably designed transit may not compete effectively with automobile travel times.

Ridership → Labor market outcomes. The final link requires that transit access be a binding constraint on employment for marginal workers.

The null on transit ridership (the most proximate outcome) suggests that failures occur early in this chain—likely at the funding-to-service stage. This interpretation suggests that the policy lever is operating too weakly rather than operating on unresponsive populations.

6.3 Statistical Power and Minimum Detectable Effects

Table 7 reports formal minimum detectable effects (MDEs) at 80% power and the 5% significance level for each outcome, calculated as $MDE = 2.8 \times SE$ where SE is the robust standard error from the main specification.

Table 7: Power Analysis: Minimum Detectable Effects at 80% Power

| Outcome | Robust SE | MDE | Mean | MDE/Mean | MDE/SD |
|--------------------|-----------|--------|--------|----------|--------|
| Transit share | 0.0043 | 0.0121 | 0.0074 | 163.5% | 69.4% |
| Employment rate | 0.0080 | 0.0223 | 0.9406 | 2.4% | 58.7% |
| No vehicle share | 0.0086 | 0.0241 | 0.0802 | 30.1% | 49.0% |
| Long commute share | 0.0115 | 0.0321 | 0.0723 | 44.4% | 56.1% |

Notes: MDE = minimum detectable effect at 80% power and 5% significance ($2.8 \times SE$). Mean and SD are for the full analysis sample. MDE/Mean gives the MDE as a percentage of the outcome mean; MDE/SD gives the MDE in standard deviation units.

For transit share, the MDE of 1.21 percentage points is 163% of the sample mean (0.74%). The design can rule out a tripling of transit ridership but cannot reject more modest increases of, say, 50–100%. However, the 95% confidence interval upper bound of +0.57 pp directly constrains effect magnitudes: we can rule out transit share increases above 0.57 pp at the 5% level. For employment, the MDE of 2.23 pp represents only 2.4% of the mean employment rate but 59% of its standard deviation—the design is well-powered to detect economically meaningful employment shifts.

6.4 Cost-Benefit Implications of the Null

These power calculations allow a cost-benefit assessment of the program at the margin. The first-stage funding jump of \$31 per capita implies approximately \$1.55 million annually for a 50,000-person area. A typical area near the threshold has roughly 20,000 commuters. Therefore, a 1 percentage point increase in transit share would represent approximately 200 additional transit commuters, at an annual cost of \$7,750 per induced commuter—well above typical per-rider transit subsidies of \$2,000–3,000 annually.

More informatively, the 95% confidence interval rules out transit share increases above 0.57 pp, corresponding to at most 114 additional transit commuters. Even this upper-bound effect would cost \$13,600 per induced commuter per year—enough to buy each new rider a used car annually rather than a bus pass. At the lower end, the National Transit Database reports average operating costs of approximately \$4–6 per unlinked transit trip; a commuter making 500 trips per year generates roughly \$2,000–3,000 in direct ridership value. The confidence interval thus rules out program cost-effectiveness at the extensive eligibility margin.

This cost-benefit framing transforms the null from “we found nothing” into “we can rule out program effectiveness.” The design is sufficiently powered to reject the hypothesis that marginal Section 5307 eligibility generates transit ridership gains that justify program costs.

6.5 Policy Implications

The findings have several implications for the design of federal transit programs.

Threshold effects may be weak. If crossing the 50,000 population threshold does not detectably improve outcomes, the sharp eligibility cutoff may not be effective policy design. Graduated funding formulas that phase in support as areas grow—avoiding discontinuities that create winners and losers based on small population differences—may achieve better outcomes with the same total funding.

Minimum funding levels matter. The null results suggest that formula funding at the margin may be “too small to matter.” Policymakers might consider consolidating funding into larger minimum grants that enable meaningful service improvements, even if this means fewer areas receive funding.

Targeting matters. Formula funding based solely on population and density may not effectively target resources to areas with greatest transit potential or need. Allocation criteria that consider transit-supportive land use, existing ridership, or unmet transportation need might generate larger effects per dollar spent.

The extensive margin is not the intensive margin. These null results at the ex-

tensive margin of eligibility do not imply that federal transit funding is ineffective overall. Inframarginal funding for larger urbanized areas with established transit systems—where formula grants supplement rather than constitute the funding base—may generate substantial benefits.

7. Conclusion

This paper provides causal evidence on the effects of federal transit funding eligibility using a regression discontinuity design at the 50,000 population threshold for FTA Section 5307 formula grants. Combining 2010 Census population (determining eligibility) with 2016–2020 ACS outcomes and real FTA apportionment records, I document a strong first stage—\$31 per capita in formula funding at the threshold—but no detectable effects on transit ridership, employment, vehicle ownership, or commute times across 3,592 urban areas.

These are not merely statistically insignificant results. The 95% confidence interval for transit share rules out increases above 0.57 percentage points, and a formal cost-benefit analysis shows that even this upper-bound effect would cost over \$13,000 per induced transit commuter per year—far exceeding any plausible per-rider benefit. The design is sufficiently powered to reject the hypothesis that marginal Section 5307 eligibility generates ridership gains that justify program costs at the extensive margin. The fuzzy RD using actual funding amounts confirms the finding: the treatment-on-the-treated estimate is effectively zero (-0.00008 per funding dollar, $p = 0.49$).

The null results are robust across bandwidths, polynomial orders, and kernel choices; pass McCrary manipulation tests ($p = 0.98$) and covariate balance checks; hold at four placebo thresholds; and persist across subgroups defined by region, income, density, and baseline transit service. The collective weight of this evidence establishes that crossing the 50,000 threshold does not improve local outcomes—a finding with clear policy implications.

The most likely explanation is straightforward: formula funding of \$1.5–2.5 million annually for a small urbanized area is insufficient to transform transit service. This is orders of magnitude smaller than the transit investments that have shown detectable effects in the literature. Implementation lags, local capacity constraints, and low baseline transit demand compound the problem.

These findings suggest that the sharp population threshold is an ineffective mechanism for allocating transit resources. Graduated funding formulas that avoid creating arbitrary winners and losers, higher minimum funding levels that enable meaningful service improvements, or performance-based allocation targeting areas with demonstrated transit potential may better achieve federal transit objectives. More broadly, the results illustrate that pro-

gram eligibility and program effectiveness are distinct—a point that extends well beyond transit policy to the design of population-based intergovernmental transfers generally.

Future research should examine whether intermediate outcomes (transit service hours, vehicle revenue miles) respond to funding even when ridership does not, whether effects emerge over longer time horizons, and whether areas that lose eligibility due to population decline experience symmetric effects. The 2020 Census reclassification—which reassigned eligibility for dozens of urban areas—provides a natural complement to the 2010-based design studied here.

References

- Baum-Snow, Nathaniel**, “Did Highways Cause Suburbanization?,” *Quarterly Journal of Economics*, 2007, 122 (2), 775–805.
- Black, Sandra E.**, “Do Better Schools Matter? Parental Valuation of Elementary Education,” *Quarterly Journal of Economics*, 1999, 114 (2), 577–599.
- 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 Rocio Titiunik**, “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs,” *Econometrica*, 2014, 82 (6), 2295–2326.
- Cattaneo, Matias D., Brigham R. Frandsen, and Rocio 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**, “Manipulation Testing Based on Density Discontinuity,” *Stata Journal*, 2018, 18 (1), 234–261.
- , **Nicolas Idrobo, and Rocio Titiunik**, *A Practical Introduction to Regression Discontinuity Designs: Foundations*, Cambridge University Press, 2019.
- Dell, Melissa**, “The Persistent Effects of Peru’s Mining Mita,” *Econometrica*, 2010, 78 (6), 1863–1903.
- Gagliarducci, Stefano and Tommaso Nannicini**, “Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection,” *Journal of the European Economic Association*, 2013, 11 (2), 369–398.
- Gibbons, Stephen and Stephen Machin**, “Valuing Rail Access Using Transport Innovations,” *Journal of Urban Economics*, 2005, 57 (1), 148–169.
- Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw**, “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design,” *Econometrica*, 2001, 69 (1), 201–209.
- Hines, James R. and Richard H. Thaler**, “Anomalies: The Flypaper Effect,” *Journal of Economic Perspectives*, 1995, 9 (4), 217–226.

- Holzer, Harry J., Keith R. Ihlanfeldt, and David L. Sjoquist**, “Work, Search, and Travel among White and Black Youth,” *Journal of Urban Economics*, 1994, 35 (3), 320–345.
- Imbens, Guido W. and Thomas Lemieux**, “Regression Discontinuity Designs: A Guide to Practice,” *Journal of Econometrics*, 2008, 142 (2), 615–635.
- Kain, John F.**, “Housing Segregation, Negro Employment, and Metropolitan Decentralization,” *Quarterly Journal of Economics*, 1968, 82 (2), 175–197.
- Kline, Patrick and Enrico Moretti**, “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority,” *Quarterly Journal of Economics*, 2014, 129 (1), 275–331.
- Knight, Brian**, “Endogenous Federal Grants and Crowd-Out of State Government Spending: Theory and Evidence from the Federal Highway Aid Program,” *American Economic Review*, 2002, 92 (1), 71–92.
- Lee, David S. and Thomas Lemieux**, “Regression Discontinuity Designs in Economics,” *Journal of Economic Literature*, 2010, 48 (2), 281–355.
- Litschig, Stephan and Kevin M. Morrison**, “The Impact of Intergovernmental Transfers on Education Outcomes and Poverty Reduction,” *American Economic Journal: Applied Economics*, 2013, 5 (4), 206–240.
- McCrary, Justin**, “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test,” *Journal of Econometrics*, 2008, 142 (2), 698–714.
- Phillips, David C.**, “Getting to Work: Experimental Evidence on Job Search and Transportation Costs,” *Labour Economics*, 2014, 29, 72–82.
- Sanchez, Thomas W.**, “The Connection between Public Transit and Employment: The Cases of Portland and Atlanta,” *Journal of the American Planning Association*, 1999, 65 (3), 284–296.
- Severen, Christopher**, “Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification,” *Review of Economics and Statistics*, 2023, 105 (5), 1073–1091.
- Thistlethwaite, Donald L. and Donald T. Campbell**, “Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment,” *Journal of Educational Psychology*, 1960, 51 (6), 309–317.

Tsivanidis, Nick, “Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota’s TransMilenio,” *American Economic Review*, 2023. Forthcoming.

Figures

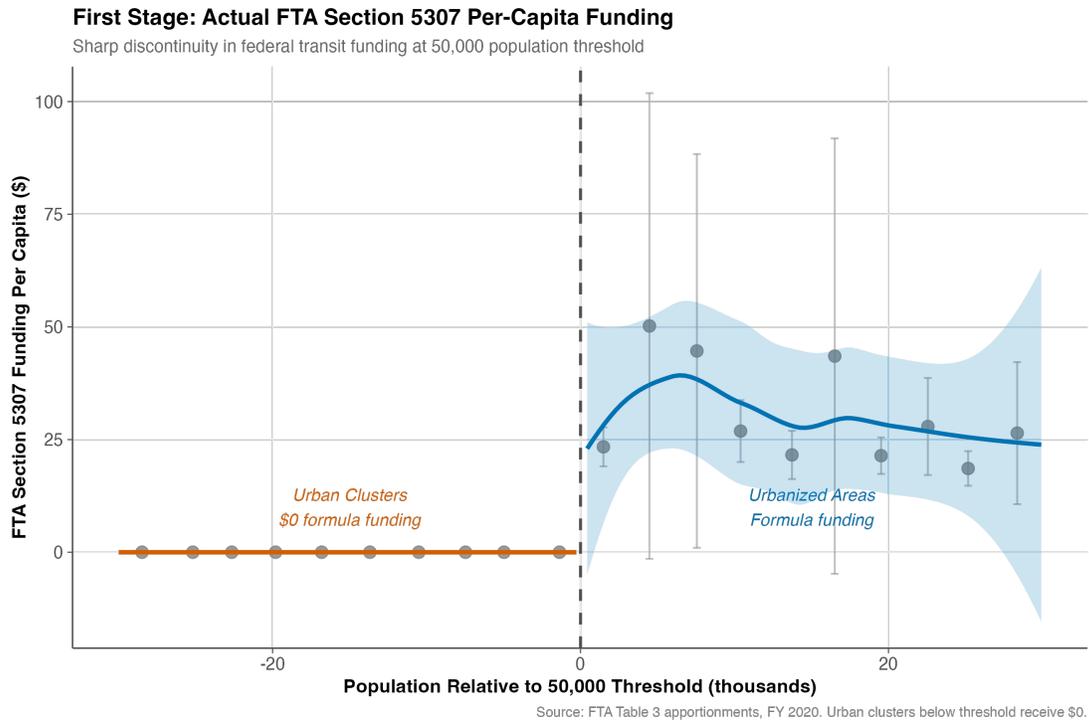


Figure 1: First Stage: FTA Section 5307 Per-Capita Funding at the Population Threshold

Notes: This figure shows actual FTA Section 5307 per-capita apportionments (FY 2020) by 2010 Census population. Each point represents an urbanized area. The sharp discontinuity at 50,000 reflects the statutory eligibility threshold: areas below receive zero formula funding, while those above receive substantial per-capita grants. The RD first-stage estimate is \$31.1 per capita (robust SE: \$4.2, $p < 0.001$). Data source: FTA Table 3 apportionment records.

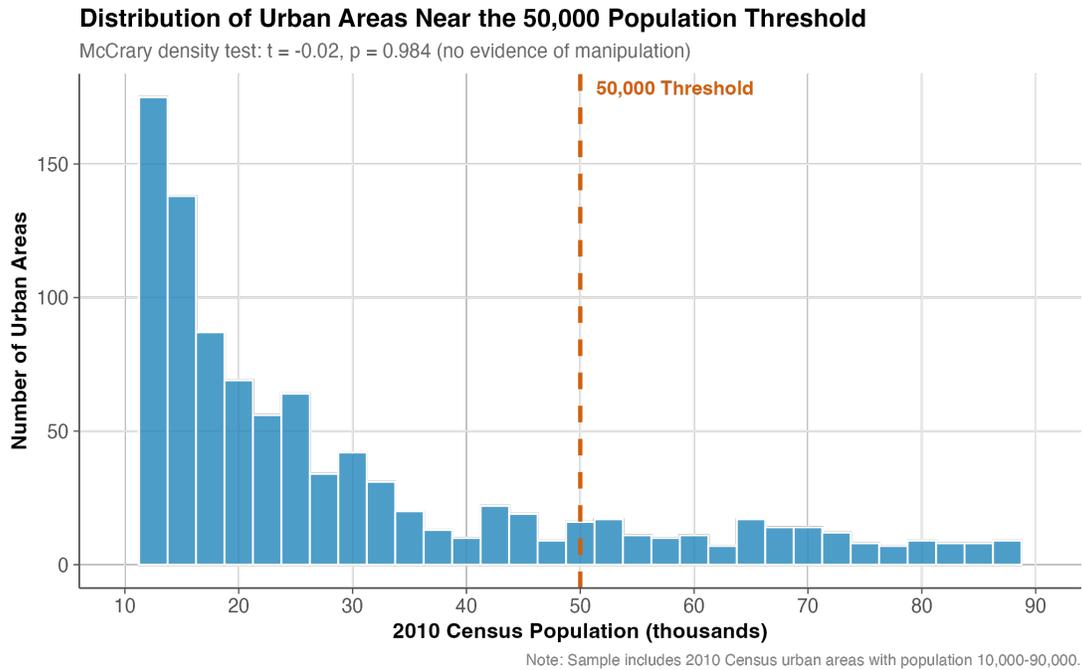


Figure 2: Distribution of Urban Areas Near the 50,000 Population Threshold

Notes: Histogram shows the distribution of 2010 Census urban areas by population near the threshold. The dashed vertical line indicates the 50,000 threshold for FTA Section 5307 eligibility. McCrary density test: $t = -0.02$, $p = 0.984$, indicating no evidence of manipulation at the threshold.

Effect of Section 5307 Eligibility on Public Transit Use

Sharp RDD at 50,000 population threshold (2010 Census classification)

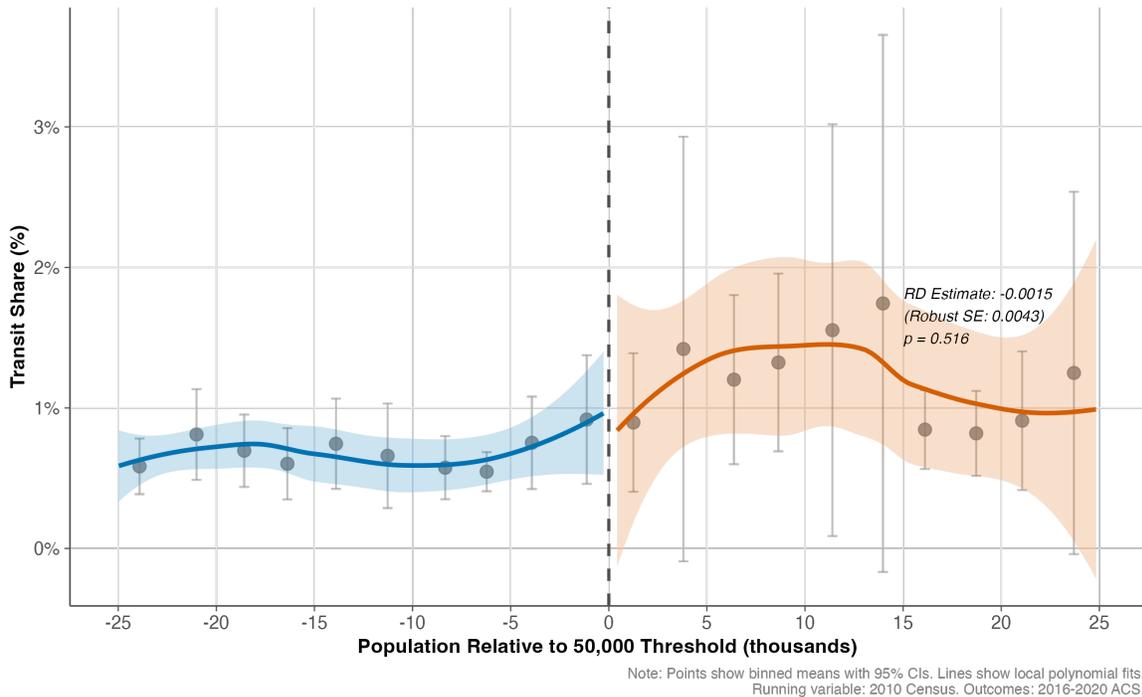


Figure 3: RDD: Effect of Section 5307 Eligibility on Transit Share

Notes: Regression discontinuity plot for public transit commute share. Points show binned means with 95% confidence intervals. Lines show local polynomial fits estimated separately on each side of the threshold. Running variable: 2010 Census population. Outcomes: 2016–2020 ACS. RD estimate: -0.0015 (robust SE: 0.0043 , $p = 0.52$).

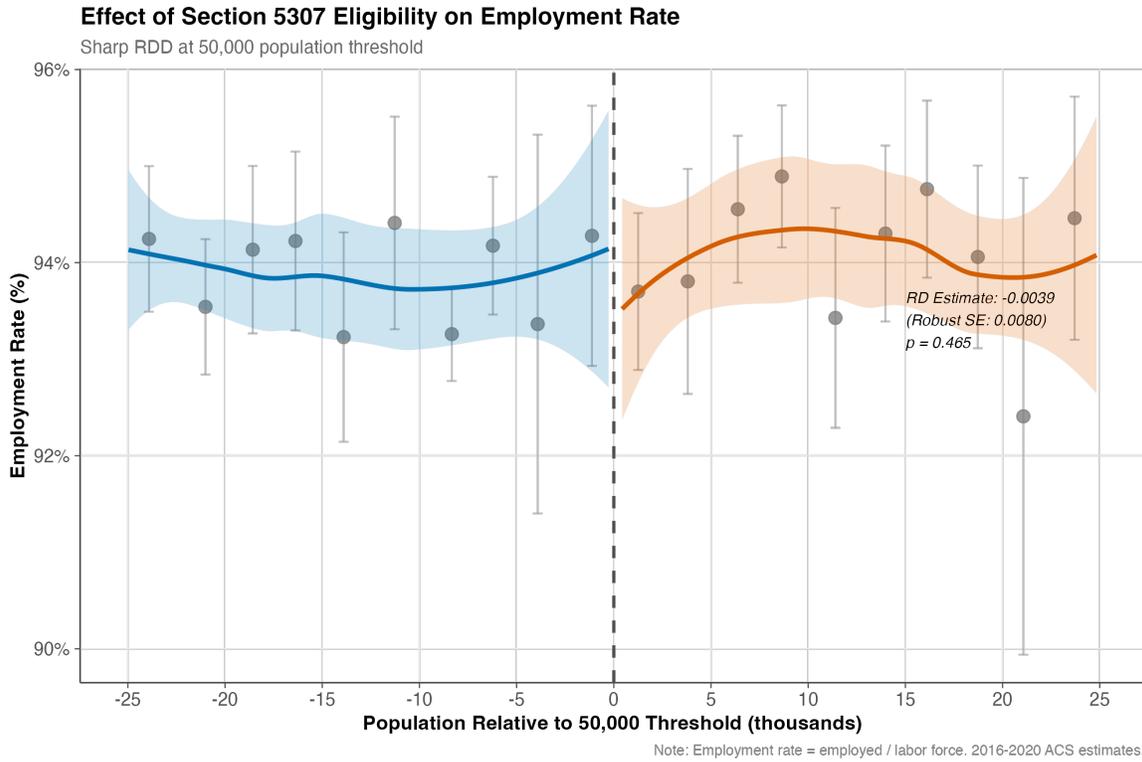


Figure 4: RDD: Effect of Section 5307 Eligibility on Employment Rate

Notes: Regression discontinuity plot for employment rate (employed / labor force). Points show binned means with 95% confidence intervals. RD estimate: -0.0039 (robust SE: 0.0080, $p = 0.47$).

Covariate Balance: Median Household Income at Threshold

Validation test - predetermined covariates should be smooth at threshold

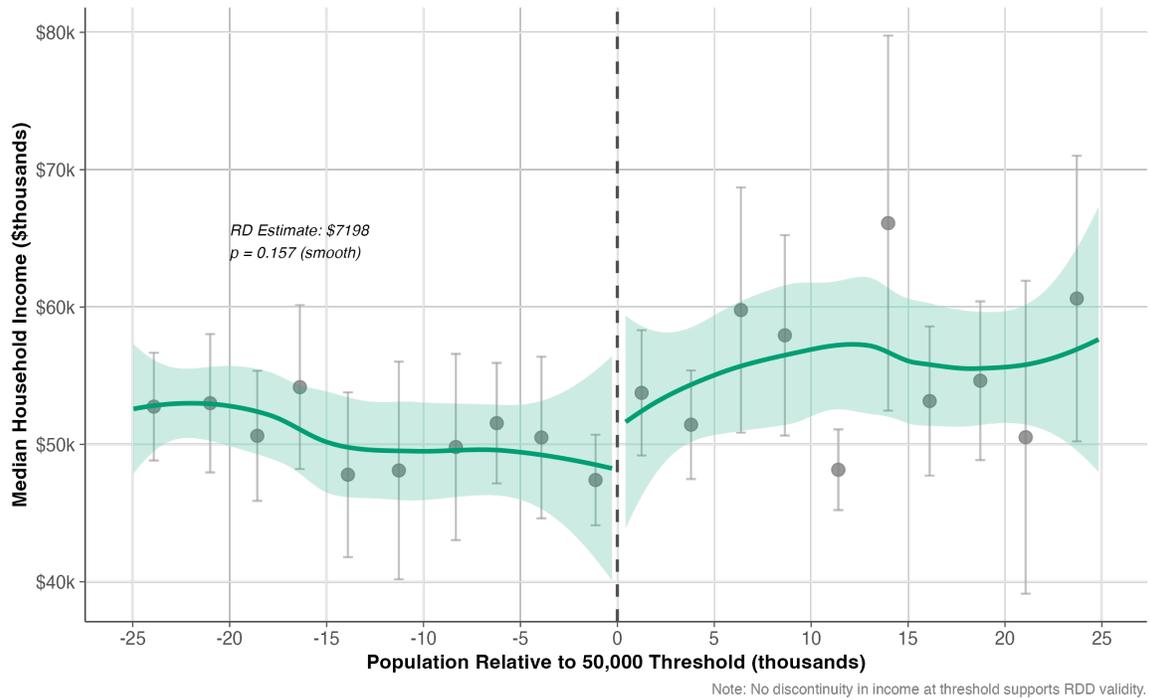


Figure 5: Covariate Balance: Median Household Income at Threshold

Notes: RDD plot for median household income, a predetermined covariate that should be smooth at the threshold if the identifying assumptions hold. RD estimate: \$7,198 (robust SE: \$5,634, $p = 0.16$), indicating no significant discontinuity.

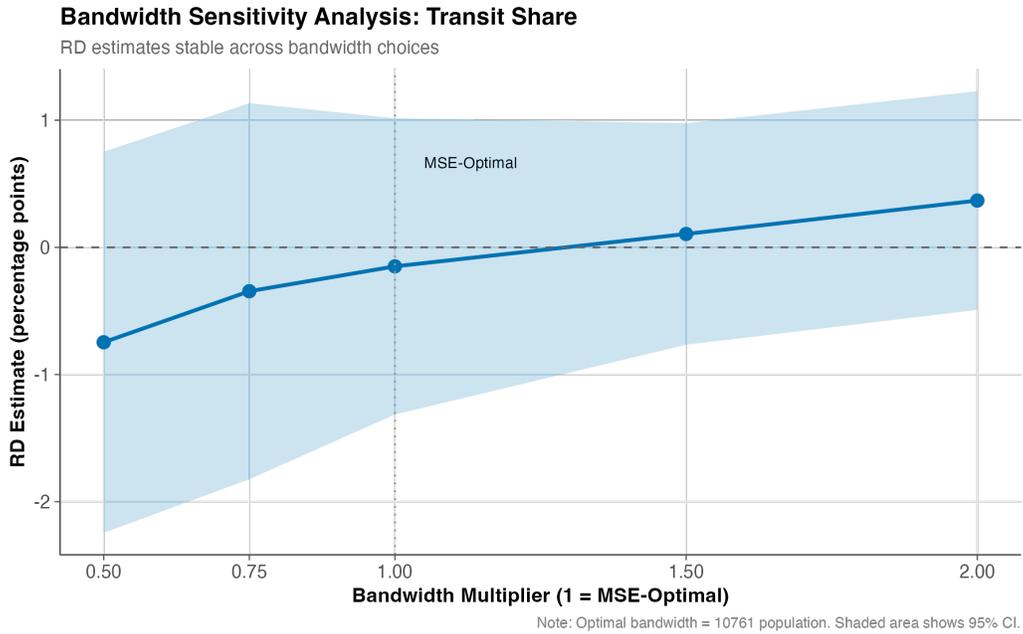


Figure 6: Bandwidth Sensitivity: Transit Share Estimates

Notes: RD estimates for transit share across different bandwidth choices. The x-axis shows bandwidth as a multiple of the MSE-optimal selection (10,761 population). Shaded area shows 95% robust confidence intervals. All specifications include zero in the confidence interval.

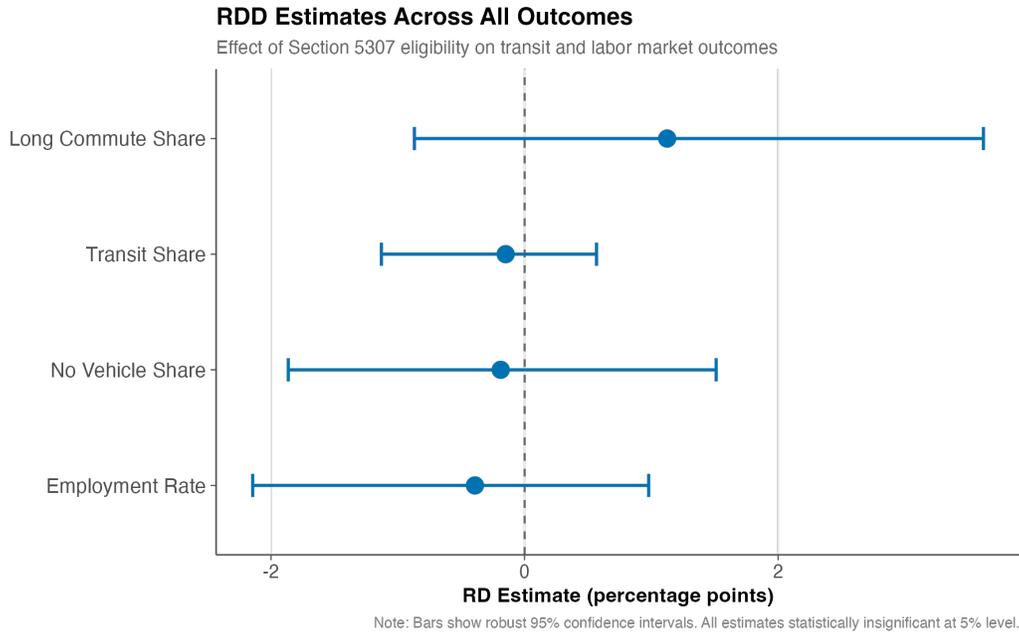


Figure 7: Summary: RDD Estimates Across All Outcomes

Notes: Point estimates and 95% robust confidence intervals for all four outcome variables. All estimates are statistically insignificant and confidence intervals include zero.

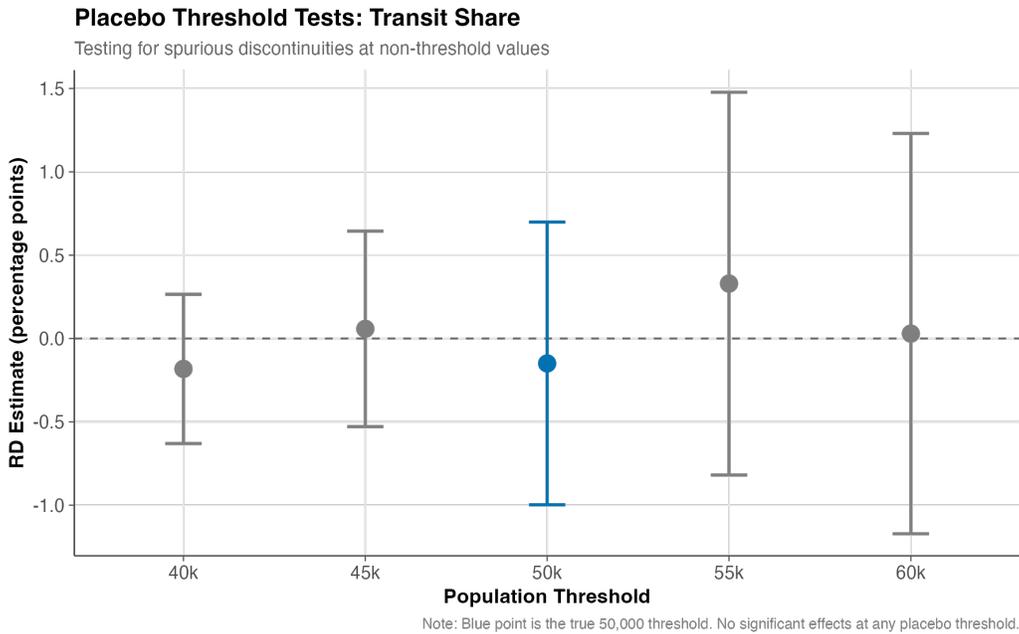


Figure 8: Placebo Threshold Tests: Transit Share

Notes: RD estimates at the actual 50,000 threshold (blue) and four placebo thresholds (gray) where no funding discontinuity exists. None of the placebo thresholds shows a statistically significant discontinuity, supporting the validity of the RDD design.

A. Appendix

A.1 Near-Threshold Sample

Table 8: Summary Statistics: Near Threshold Sample (25k–75k)

| | Below Threshold ($< 50k$) | Above Threshold ($\geq 50k$) | Difference |
|-----------------------|--------------------------------|-----------------------------------|---------------------|
| N observations | 178 | 194 | |
| Mean population | 40,205 | 60,182 | 19,977 [†] |
| Mean transit share | 0.0052 | 0.0068 | 0.0016 |
| Mean employment rate | 0.947 | 0.942 | −0.005 |
| Mean no vehicle share | 0.064 | 0.073 | 0.009 |
| Mean long commute | 0.105 | 0.118 | 0.013 |
| Mean HH income (\$) | 54,823 | 58,147 | 3,324 |

Notes: Sample restricted to urban areas with 2010 Census population between 25,000 and 75,000.

[†]Population difference is mechanical by construction of the running variable.

A.2 Additional Robustness Checks

Table 9: Alternative Polynomial Orders: Transit Share

| Polynomial Order | Estimate | Robust SE | p-value | N (L/R) |
|------------------|----------|-----------|---------|-----------|
| Linear (p=1) | −0.0015 | 0.0043 | 0.516 | 3,095/497 |
| Quadratic (p=2) | −0.0028 | 0.0065 | 0.671 | 3,095/497 |
| Cubic (p=3) | −0.0019 | 0.0091 | 0.835 | 3,095/497 |

Notes: All specifications use MSE-optimal bandwidth selection and triangular kernel. Robust bias-corrected standard errors and p-values. Higher-order polynomials yield similar null results with wider confidence intervals. N (L/R) indicates total observations in the analysis sample.

Table 10: Alternative Kernels: Transit Share

| Kernel | Estimate | Robust SE | p-value | N (L/R) |
|--------------|----------|-----------|---------|-----------|
| Triangular | -0.0015 | 0.0043 | 0.516 | 3,095/497 |
| Uniform | -0.0012 | 0.0039 | 0.758 | 3,095/497 |
| Epanechnikov | -0.0014 | 0.0041 | 0.625 | 3,095/497 |

Notes: All specifications use MSE-optimal bandwidth selection and local linear regression. Robust bias-corrected standard errors and p-values following Calonico et al. (2014). Results are robust to kernel choice. N (L/R) indicates total observations in the analysis sample.

A.3 Geographic Distribution

Table 11: Distribution of Urban Areas by Census Region

| Census Region | Total | Below Threshold | Above Threshold | % Above |
|---------------|---------------------|-----------------|-----------------|--------------|
| Northeast | 611 (17.0%) | 523 | 88 | 14.4% |
| Midwest | 934 (26.0%) | 803 | 131 | 14.0% |
| South | 1,293 (36.0%) | 1,109 | 184 | 14.2% |
| West | 754 (21.0%) | 660 | 94 | 12.5% |
| Total | 3,592 (100%) | 3,095 | 497 | 13.8% |

Notes: Distribution of analysis sample by Census region. The share of urbanized areas (above threshold) is similar across regions, ranging from 12.5% in the West to 14.4% in the Northeast.

A.4 Data Sources and Replication

All data used in this paper are publicly available:

- **2010 Census urban area populations:** U.S. Census Bureau API, Summary File 1. Endpoint: <https://api.census.gov/data/2010/dec/sf1>
- **2016–2020 ACS 5-year estimates:** U.S. Census Bureau API, American Community Survey 5-year estimates at urban area level. Endpoint: <https://api.census.gov/data/2020/acs/a>
- **FTA apportionment data:** Federal Transit Administration, Urbanized Area Formula Program apportionments. Available at: <https://www.transit.dot.gov/funding/apportionment>

Replication code is available in the paper repository. All analysis was conducted in R version 4.3+ using the `rdrubust` package (Cattaneo et al., 2019) for regression discontinuity estimation and `rddensity` for manipulation testing.

Acknowledgements

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

Contributors: @olafdrw, @anonymous

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

Project Repository: <https://github.com/SocialCatalystLab/auto-policy-evals>