Mandating Recycling: Evidence from Pennsylvania's Density Cutoff

Costanza Biavaschi*

June 30, 2026

Preliminary and incomplete. Please do not circulate without the author's permission.

Abstract

This paper estimates the long-run effect of assigning municipalities to a statutory recycling mandate. Pennsylvania's Act 101 requires municipalities with 5,000–10,000 residents to adopt recycling ordinances, provide curbside and leaf-waste collection, impose commercial source-separation duties, and report tonnage when population density exceeds 300 persons per square mile. I exploit this inherited statutory threshold in a regression-discontinuity design using municipal administrative records from 2001–2024. Mandate assignment raises administratively recorded total recycling per capita by 99.3 percent; residential and commercial recycling rise by 53.1 and 147.2 percent. A decomposition shows that the effect is not driven by municipalities newly entering the reporting system. Residential gains mainly reflect the broader compliance bundle rather than large additional household sorting among already-served municipalities. Commercial gains are concentrated in covered materials already reported on both sides of the cutoff. Municipal mandates raise recorded recycling primarily by reshaping the administrative system through which generators separate covered materials, municipalities document tonnage, and the state monitors compliance.

JEL Codes: Q53, Q58, H73, K32, C21

Keywords: recycling, mandates, regression discontinuity, partial identification, extended producer responsibility

*Department of Economics and Management, University of Pisa. I gratefully acknowledge the Pennsylvania Department of Environmental Protection (DEP) and John Nantz in particular for generously sharing data and offering valuable insights on the program. The views expressed here are my own and do not represent the DEP's position. Any errors are entirely my responsibility.

1 Introduction

Most causal evidence on recycling policy studies household-facing margins: curbside access, unit pricing, deposit-refund schemes, and information treatments. Yet the recycling system that modern policy tries to govern is broader than household sorting. Roughly 85 percent of U.S. recycled municipal solid waste tonnage comes from outside the single-family curbside stream (National Academies of Sciences, Engineering, and Medicine, 2025), and municipal recycling statutes often operate through local ordinances, commercial-generator obligations, reporting systems, and compliance monitoring. This question is increasingly relevant because Extended Producer Responsibility (EPR) statutes emphasize producer financing and accountability (Walls, 2011, Joltreau, 2022, OECD, 2024), but their performance still depends on local systems that determine collection coverage, material scope, reporting, and compliance. This creates a basic empirical question: does assigning a municipality to a statutory recycling obligation raise municipal recycling tonnage, and through which administrative and sectoral margins?

The closest causal evidence observes policy variation at other levels. Viscusi et al. (2023) link within-household changes in self-reported recycling to changes in local and state recycling conditions, including moves into states with deposit laws or effective recycling laws. Ogieriakhi and Wang (2024) and Anglou et al. (2024) study state-level commercial recycling or food-waste mandates using aggregate waste-disposal outcomes and counterfactual trends across jurisdictions. This evidence is informative about policy regimes, but it identifies effects from household moves or cross-jurisdiction policy contrasts, where recycling laws are bundled with other state, administrative, and market differences, rather than from assignment of a municipal mandate within a common policy regime.

I study that assignment margin using Pennsylvania's Act 101 of 1988. The setting provides, to my knowledge, the first regression-discontinuity estimate of a statutory municipal recycling mandate and identifies its long-run effect. Among municipalities with 5,000–10,000 residents, Act 101 assigns the obligation when population density exceeds 300 persons per square mile. Mandated municipalities must adopt a recycling ordinance, provide curbside collection for designated household materials, collect leaf waste, conduct public education, impose source-separation duties on commercial, municipal, and institutional generators, and report annual recycling quantities through the statewide Act 101 system. I exploit the 300 ppsm rule in a regression-discontinuity design comparing municipalities just above and below the cutoff within the same statewide legal, grant, reporting, and recycling-market environment.

The identifying variation comes from a threshold written for a different purpose. It entered Pennsylvania law in an 1899 township-classification act, when statewide density was 141 persons per square mile, and Act 101 reused the same number nearly ninety years later. I find no bunching at the threshold and no discontinuities in predetermined municipal characteristics. I also verify that other pre-existing Pennsylvania statutes citing 300 persons per square mile do not create binding discontinuous obligations within the 5,000–10,000 population band. The design therefore isolates mandate assignment within a common policy regime, avoiding the compound-treatment problem that often arises in population-threshold designs (Grembi et al., 2016, Eggers et al., 2018).

The mandate nearly doubles administratively recorded recycling. At the cutoff, assignment to Act 101 raises per capita total recycling recorded by the Pennsylvania Department of Environmental Protection (DEP) by 99.3 percent ($\hat{\tau} = 0.690$, s.e. $= 0.195$). Evaluated at the below-cutoff mean of 200.0 pounds per resident per year, the estimate implies an above-cutoff rate of 399 pounds. For scale, this is close to the U.S. per capita recycling benchmark of 423 pounds (U.S. Environmental Protection Agency, 2020). The response appears in both streams: residential recycling rises by

53.1 percent and commercial recycling by 147.2 percent.

These estimates are effects on the administrative object Act 101 uses for compliance monitoring, grant payments, and public accounting. The central interpretation question is therefore what the discontinuity in DEP-recorded tonnage represents. One possibility is reporting entry: municipalities above the cutoff may newly appear in individual DEP records. Institutionally, this explanation is unlikely to drive the result: the cutoff does not change grant payments, since municipalities on both sides can report eligible tonnage and receive performance grants through the same statewide system. Consistent with this structure, individual municipal reporting is high and smooth at the threshold (Section 3).

To characterize what the discontinuity represents, I decompose the recorded-tonnage response into three administrative margins: *activation*, whether a municipality records any positive tonnage; *composition*, which material groups appear in the records; and *intensification*, tonnage within already-active material groups. Activation explains little of the total effect. The mandate does broaden recorded material scope, especially through leaf waste and other newly recorded categories. But the largest total and commercial contributions come from intensification within material groups that already appear in municipal records, with roughly half of the total and commercial discontinuities coming from non-leaf materials covered by Act 101.

This decomposition disciplines the household interpretation. In the full sample, residential recycling rises substantially. But when I restrict the comparison to municipalities with curbside service already recorded at the start of the panel, covered residential intensification essentially disappears. The residential effect is therefore best read as the long-run effect of the broader Act 101 residential compliance bundle—curbside program provision, leaf-waste collection, covered-material scope, public education, and reporting formalization—rather than as evidence that the mandate substantially increased household sorting intensity once curbside service was already in place.

The commercial response points to a different margin. The commercial discontinuity remains large in the curbside subsample, and covered-material intensification accounts for roughly half of it. Act 101 does not require municipalities to collect commercial recycling directly. Instead, it requires commercial, municipal, and institutional generators to separate covered materials and requires municipalities to document that tonnage within the Act 101 reporting system. Because activation is small and material-group composition does not account for the main commercial response, the estimate does not reflect municipalities newly appearing with positive commercial tonnage or newly adding commercial categories to their reports. What the mandate does is raise tonnage within covered commercial materials already reported on both sides of the cutoff. That increase may reflect additional source separation by generators newly facing municipal source-separation obligations, more complete documentation of private-hauler recycling within existing reporting channels, or both. I read the commercial estimate as covered commercial intensification within the municipal compliance system.

The long-run horizon is central to this interpretation. Act 101 was enacted in 1988, while the administrative panel spans 2001–2024. The estimates therefore capture steady-state differences after initial program build-out, enforcement routines, and reporting practices had stabilized. Figure 1 shows persistent treated–control gaps throughout the panel, and decade-specific estimates are similar across the commodity-market regimes spanned by the data, including the post-2018 National Sword disruption to mixed-paper and plastics export markets (Sigman and Strow, 2024). This is the relevant horizon for EPR and related institutional policies, which are designed to operate over decades rather than adoption windows.

The paper makes three contributions. First, it identifies the within-regime effect of municipal mandate assignment. Second, it clarifies how to interpret administrative recycling outcomes

by decomposing the response into activation, composition and intensification, adapting scale–composition–technique logic (Shapiro and Walker, 2018, Greenstone and Hanna, 2014) to an administrative count setting in which measurement and compliance are part of the policy itself. Third, the DEP records separately report residential and commercial tonnage, allowing the design to estimate a mandate effect on recorded commercial recycling. That commercial margin is largely absent from the quasi-experimental recycling literature and is central to the institutional systems through which modern recycling policy is implemented. Administrative tonnage is not merely a measurement device; it is the object through which recycling mandates monitor compliance, allocate grants, and make local recycling activity visible.

2 Institutional Background

Pennsylvania's Municipal Waste Planning, Recycling and Waste Reduction Act (Act 101 of 1988) assigns recycling obligations to individual municipalities through a population-and-density rule. A municipality is mandated if its population is at least 10,000, or if its population is greater than 5,000 but less than 10,000 and its density is greater than 300 persons per square mile (ppsm). Mandated municipalities must establish a curbside recycling program, collect designated materials, conduct public education, and report annual recycling quantities through the Act's county-based reporting system. The 300 ppsm line within the greater-than-5,000 and less-than-10,000 population band is the source of within-state variation this paper exploits.

This section documents the institutional facts that matter for the empirical design. First, the 300 ppsm threshold predates recycling policy by nearly a century. Second, other Pennsylvania statutes that reference the same line do not create a discontinuous recycling-relevant policy change within the analysis sample. Third, both sides of the threshold operate inside the same statewide recycling-finance and reporting system; what changes at the cutoff is the statutory obligation to operate a compliant municipal program.

Origins of Act 101. Recycling policy in the United States is decentralized. The federal government set a broad framework for solid waste management under the Resource Conservation and Recovery Act, but municipal waste planning and household recycling remain primarily state and local responsibilities. The expansion of recycling in the late twentieth century therefore proceeded through a wave of state laws rather than a single national mandate.

Pennsylvania enacted the Municipal Waste Planning, Recycling and Waste Reduction Act (Act 101) in 1988, amid acute disposal pressure. Statewide disposal capacity had fallen to 12–18 months, while more than 1,100 substandard unlined landfills and town dumps were closing for environmental inadequacy (Seif, 1997). Against this backdrop, Act 101 was unusually stringent among late-1980s state recycling laws: by 1990, Pennsylvania was one of only five states, plus the District of Columbia, requiring household source separation as the operational standard (National Solid Wastes Management Association, 1991, Viscusi et al., 2013).

The Act combined municipal recycling mandates with county waste planning and dedicated recycling finance.[1] Among source-separation states, Pennsylvania was unusual in assigning the operational mandate to individual municipalities through a population–density threshold rather than applying it universally (National Solid Wastes Management Association, 1991).

[1]Pennsylvania Municipal Waste Planning, Recycling and Waste Reduction Act, Act of Jul. 28, 1988, P.L. 556, No. 101 (53 P.S. §§ 4000.101 et seq.). The municipal mandate is established in § 1501; the tipping fee in § 701; planning, implementation, and performance grants in §§ 901–904; and penalties in § 1704.

The 300 ppsm threshold. The 300 ppsm threshold predates modern recycling policy. It first appeared in Pennsylvania's 1899 township-classification act, which divided townships into first and second classes by population density and placed townships with at least 300 inhabitants per square mile in the first class. The classification was designed to give rapidly urbanizing townships near Philadelphia and Pittsburgh a stronger local-government form for managing suburban growth. The same numerical line was subsequently reused in municipal waste planning and then in Act 101's recycling mandate.[2]

The line was therefore drawn for municipal classification, not recycling economics. In 1900, statewide density was 141 ppsm, and only a small set of townships around Philadelphia and Pittsburgh reached the threshold; most dense settlements were already governed as cities or boroughs under separate codes. By 1980, statewide density had reached roughly 265 ppsm, and within the 5,000–10,000 population band the inherited line assigned the recycling mandate mainly to municipalities in the Philadelphia metropolitan area, the Pittsburgh core, and the Lehigh Valley rather than to the rural majority.[3]

The initial Act 101 mandate lists used the 1980 decennial Census, the most recent available at enactment. Municipalities with populations of at least 10,000 were required to comply by September 1990; municipalities with more than 5,000 but fewer than 10,000 residents and density greater than 300 ppsm were required to comply by September 1991. Municipalities crossing either threshold after a later decennial Census must establish a compliant program within two years of the census data becoming official.

Co-located statutes and identification. Three other Pennsylvania statutes also reference the 300 ppsm line. None creates a discontinuous recycling-relevant policy change within the 5,000–10,000 population band.

First, the same threshold makes townships eligible for first-class status, which could independently affect municipal service capacity. Eligibility is not automatic, however: a township above 300 ppsm becomes first class only after local voters approve reclassification in a referendum. This additional political step breaks any mechanical link between the density cutoff and township class. Consistent with that institutional structure, first-class status does not jump at 300 ppsm in the analysis sample; it is essentially absent below 250 ppsm and rises smoothly only well above the threshold (Table 1; see Section 4.1). Among the ninety-three current first-class townships statewide, ninety-one had already attained first-class status when Act 101 was enacted in 1988.[4] Excluding these municipalities leaves the estimates stable (see Section 7).

Second, Act 97 of 1980 used the 300 ppsm line for municipal waste planning. That requirement was procedural and disposal-oriented: municipalities above the line had to submit a written plan

[2]The 1899 statute is the Act of Apr. 28, 1899, P.L. 104, No. 86. The threshold was later codified in Act 331 of June 24, 1931 (53 P.S. § 55101 et seq.), the First Class Township Code, and reaffirmed in the Second Class Township Code, 53 P.S. § 65201. Act 97 of 1980 used the same threshold for municipal waste planning.

[3]1900 density: 6,302,115 residents over 44,817 square miles. 1980 density: 11,863,895 residents over 44,743 square miles (U.S. Census).

[4]The most recent pre-enactment transition occurred in 1976, twelve years before the mandate. Classification status is from the Pennsylvania Department of Community and Economic Development; transition dates are from the Local Geohistory Project, cross-verified against the PSATC member roster. Two townships in the broader panel reclassified from second to first class during the sample period (Lower Macungie in 2008, Pocono in 2014), but neither contributes to the RD estimates at the time of reclassification: Lower Macungie's population (19,220 in the 2000 cohort, 30,633 in 2010) places it outside the 5,000–10,000 band in every year, and Pocono exits the band in 2010 (population 11,065), four years before its reclassification. The choice of how to code first-class status over time is therefore immaterial for the analysis sample.

to DEP covering waste-system extension, financing, and timing. The statute permitted but did not require recycling components in a plan, listed no source-separation mandate, recycling-rate target, curbside requirement, or materials list, and provided no operating funds for recycling. The recycling architecture — recycling ordinance, materials list, $2-per-ton disposal fee, Recycling Fund, and tonnage-based performance grants — arrived only with Act 101 in 1988. Act 101 superseded this earlier planning regime by moving primary waste-planning responsibility to counties. Any residual effect of Act 97 would therefore have to operate through planning capacity accumulated under a 1980–1988 procedural requirement, not through a contemporaneous recycling treatment at the cutoff.

Third, the Municipal Authorities Act bars opt-out from solid-waste-authority service for property owners in areas above 300 ppsm, but only in 8th-class counties. Pennsylvania's 8th-class counties (Cameron, Forest, Fulton, Montour, Potter, and Sullivan) contain no municipality meeting the 5,000-population floor of the analysis sample, so this provision cannot affect any observation used in the RD.

Compliance requirements and enforcement. Act 101 made municipal recycling a documented compliance obligation rather than only a curbside-access rule. A mandated municipality must adopt a recycling ordinance, provide regular curbside collection for designated household materials, collect leaf waste, conduct public education, impose recycling duties on commercial, municipal, and institutional generators, and report annual recycling quantities through the Act's county-based reporting system. These requirements moved mandated municipalities beyond the voluntary drop-off centers and county-organized programs that characterized much pre-mandate recycling activity.[5]

Those local duties operate inside statewide finance and reporting infrastructure. A $2-per-ton disposal fee funds county waste planning, recycling coordinators, municipal implementation grants, and annual performance grants. Counties prepare ten-year municipal-waste plans and house recycling coordinators, so mandated and voluntary municipalities within the same county share a common planning and administrative environment. The county fixed effects introduced in Section 4 absorb cross-county differences in that infrastructure.

This institutional structure matters for interpreting the cutoff. The 300 ppsm threshold does not create access to Act 101's reporting-and-grant system; municipalities on both sides can report grant-eligible tonnage and receive performance grants. What changes at the cutoff is legal obligation. Above the threshold, participation in the same statewide system becomes part of a statutory compliance bundle covering designated materials, public education, commercial generators, mandatory reporting, DEP enforcement authority, and civil penalties.[6]

[5]For households, §1501(c) requires curbside collection of at least three materials from a list comprising clear glass, colored glass, aluminum cans, steel and bimetallic cans, high-grade office paper, newsprint, corrugated paper, and plastics. Leaf-waste collection is required in addition. Commercial, municipal, and institutional generators within mandated municipalities face a separate requirement under §1501(c)(iii), covering high-grade office paper, corrugated paper, aluminum cans, and leaf waste.

[6]Implementation grants are authorized under § 902, performance grants under § 904, and civil penalties under § 1704. § 902(a) reimburses approved startup costs for both mandated and voluntary adopters; § 902(d) equalizes grant priority between voluntary mandatory-source-separation programs and § 1501-mandated municipalities. § 904 performance grants are available to any reporting municipality, but mandated municipalities and non-mandated recipients of more than $10,000 in a year must meet seven program requirements to spend the funds without restriction (§ 904(d)(4)–(6)). Per-ton payment is set in § 904(b), tied to the source-separated materials enumerated in § 1501(a); § 904(d)(3) excludes tonnage produced by leaf or other organics composting facilities. § 1704 authorizes civil penalties up to $10,000 per violation, with each day of continued violation treated as a separate offense.

This statewide system also shapes how the tonnage outcome should be read. For mandated and voluntary municipalities alike, any performance-grant claim must be backed by material-specific reported tons that are grant-relevant and subject to documentation and review. Unsupported tons can be disallowed, and DEP may restrict or deny disbursement when grant prerequisites are not met. The per-ton grant therefore creates an incentive to report grant-eligible recycling, and potentially to overstate it, but both the incentive and the documentary constraints operate through the statewide system rather than appearing only above 300 ppsm. The tonnage measure is not a voluntary survey response. It is the administrative quantity through which Act 101 pays grants, monitors compliance, and accounts publicly for municipal recycling.

Mandated municipalities in the 5,000–10,000 population band were required to establish compliant programs by September 1991, a decade before the panel begins. The 2001–2024 panel therefore measures the long-run equilibrium effect of mandatory compliance within Act 101's statewide reporting-and-grant system. The next section describes the DEP administrative records that measure this tonnage outcome. Section 6 returns to whether the discontinuity reflects activation, material-group composition, or intensification within already-active material groups.

3 Data

Data Sources and Variables. The analysis combines three sources: DEP Act 101 administrative reports, the Census Gazetteer, and 1980 NHGIS demographic files. The post-2001 panel is built from the Gazetteer universe of Pennsylvania municipalities observed annually from 2001 through 2024; pre-mandate covariates are available for the 2,572 minor civil divisions with stable 1980–1990 boundaries.[7]

The DEP records measure recycling tonnage in Pennsylvania's Act 101 reporting system, the administrative quantity used for compliance monitoring, performance grants, and public accounting. Because these filings enter per-ton grant payments and are subject to DEP review, they are the statute's operating measure of municipal recycling rather than a voluntary survey. I construct three outcomes per municipality-year: total, residential, and commercial recycling per capita, observed across the full 2001–2024 panel.[8]

To my knowledge, the DEP records are the first municipal panel to pair residential and commercial recycling tonnage for more than two decades. This directly addresses the measurement gap emphasized by National Academies of Sciences, Engineering, and Medicine (2025): national recycling measures rely heavily on modeled aggregates, household surveys, residential program data, or sectorally aggregated administrative panels, leaving direct observations of commercial behavior scarce.[9] The sectoral split is central to the decomposition in Section 6: residential tonnage mea-

[7]Counts of Pennsylvania municipalities vary slightly across sources and reporting tables. The 2,572 figure refers to MCDs with stable boundaries between the 1980 and 1990 NHGIS Censuses, the universe on which the pre-mandate covariates are constructed. Table A1 reports 2,587 unique municipalities in the 2001–2024 analysis universe; the additional fifteen reflect boundary-change entities arising under the 1994 Municipal Consolidation Act (Pennsylvania Local Government Commission, 2025) that received new Gazetteer identifiers after 1990. The estimation sample is anchored on the post-2001 Gazetteer roster.

[8]In the empirical sections, I use "recycling" as shorthand for this recorded tonnage measure.

[9]Long municipal panels are typically residential-only, including Allers and Hoeben (2010) for Dutch municipalities, Best and Kneip (2019) for German household-level data, and Carattini et al. (2018) for Swiss pricing-by-the-bag, or report aggregate disposal tonnage without a sectoral split, as in Akbulut-Yuksel and Boulatoff (2021). Where commercial tonnage is observed separately, the panel is short and not paired with a matched residential series (Ogieriakhi and Wang, 2024). Cross-sectional U.S. work likewise focuses on the residential side (Jenkins et al., 2003).

sures household-sourced material, while commercial tonnage measures material from commercial, municipal, and institutional establishments.

This split is not inferred ex post: municipalities report residential and commercial tonnage on DEP's standardized §904 Residential & Commercial Tonnage Summary Forms, entered by hauler and material, and DEP permits estimated allocations only when the basis is documented and weight-slip records are retained for audit. Because approved commercial tonnage is capped at residential tonnage for the main Base and Bonus Awards, the grant formula also limits the incentive to mechanically shift tonnage between the two streams.[10] The records also report material-category breakdowns, distinguishing covered materials from non-covered materials. A separate DEP curbside-program registry identifies which municipalities operate municipal curbside collection in each year. The 2001 curbside coverage measure is used in the heterogeneity analysis in Section 6.

Treatment assignment is constructed from Act 101's legal rule. The Census Gazetteer defines the universe of Pennsylvania municipalities and supplies population and land area, from which I construct the density running variable. I merge the DEP records to the Gazetteer on municipality name and county identifiers, matching 99.9% of non-countywide records (24 of 36,953 fail exact-name and manual lookup).[11] Act 101 assigns treatment using the decennial benchmark in force at the time: the 2000 benchmark governs 2001–2009, the 2010 benchmark governs 2010–2019, and the 2020 benchmark governs 2020–2024. Per capita outcomes use the same Census-benchmark population in the denominator, so running variable and outcome scale update jointly at Census decades rather than interpolating annually.

Controls are predetermined with respect to the mandate, the running variable, and recycling outcomes. They are drawn from the 1980 decennial Census via the National Historical Geographic Information System (NHGIS) — the last benchmark before Act 101's 1988 enactment.[12]

The control vector contains twelve covariates, chosen ex ante to span the four channels through which Act 101's legislative history flags plausible confounding. *Demographic composition* (white share, prime-age share) addresses concerns of a density gradient in pro-environmental preferences as the source of the discontinuity. *Fiscal capacity* (college-educated share, median household income, unemployment rate) addresses the concern that denser municipalities have greater resources to fund startup costs and sustain ongoing compliance. *Housing structure* (owner-occupancy share, single-family detached share) captures density gradients in the physical feasibility of curbside collection— truck access, container storage, single-point pickup. *Industrial composition* (1980 employment shares in manufacturing; wholesale and retail trade; FIRE and business services; personal, health, and educational services; and public administration) addresses sectoral variation in commercial-waste generation as a potential source of the commercial-stream discontinuity. The seven demographic, fiscal-capacity, and housing-structure characteristics follow Jenkins et al.

[10]By rule, multi-family dwellings are classified as residential regardless of hauler route type, and commercial includes institutional and government generators. DEP conducts annual documentation reviews of a sample of applicants, with independent audit authority retained by the Office of the Auditor General and the Office of the Treasurer for four years from submission; selected municipalities must produce hauler and end-market weight slips and a list of the commercial establishments whose tonnage was counted, and any tonnage not substantiated at audit is disallowed and the grant award adjusted. Reported tonnages for audited municipality-years are therefore post-review approved figures rather than raw self-reports.

[11]Countywide entries are aggregated DEP records that span an entire county rather than a single minor civil division and are excluded from municipality-level matching.

[12]Contemporaneous demographics reflect sorting induced by the mandate or by correlated municipal policies, introducing post-treatment bias. The 1980 running variable used in the balance check (Table 1) precedes any version of the assignment rule.

(2003), Beatty et al. (2007), and Allers and Hoeben (2010); the five industry shares are added here to discipline the commercial margin. All twelve are listed in Table 1. Median household income is BLS CPI-U deflated to March 2026 dollars.

To construct pre-mandate density, I pair 1980 NHGIS population with 1990 Gazetteer land area, since the 1980 Gazetteer is unavailable. Pennsylvania municipal boundaries were stable across the 1980s: the 1980 and 1990 Censuses record the same 2,572 MCDs.[13]

Sample Construction. Table A1 summarizes the construction of the estimation sample, from the full Gazetteer universe of Pennsylvania municipality-years through the population-band, individual-reporting, outlier, and bandwidth restrictions described below.

The 5,000–10,000 population band is the range in which the 300 ppsm density rule is the binding assignment margin (Section 2). Within this band, Act 101 assigns the mandate at a sharp density cutoff, generating the regression discontinuity formalized in Section 4; outside it, the mandate is determined by population alone and density is irrelevant to treatment status. Starting from the full Gazetteer universe, this restriction yields 373 unique municipalities observed over the 2001–2024 panel, for a total of 7,521 municipality-years. Coverage is not perfectly balanced across years because population benchmarks change at Census decades: municipalities enter and leave the 5,000–10,000 range as population shifts.[14]

Because the recycling outcomes are constructed from municipality-level DEP filings, I first document the coverage of individual tonnage reports before imposing the outcome-sample restriction. Reporting coverage is very high within the 5,000–10,000 population band. Of the 7,521 municipality-years, 6,994 (93 percent) carry individual recycling tonnage reports. Most of the remaining cases are still covered by DEP records: 491 municipality-years (6.5 percent) appear in county-wide filings, while just 36 (0.5 percent) have no DEP record at all. The non-individual share is concentrated below the threshold (404 of 491 county-aggregate cases), consistent with county-wide filing conventions in less dense areas rather than selective non-reporting by mandated municipalities. Non-individual status is year-specific rather than a fixed municipal trait: only 2 of 373 municipalities are never individually observed, and 264 report individually in every year they appear in the 5,000–10,000 population band. Taken together, the descriptive evidence points to near-universal DEP coverage with some county-level aggregation below the cutoff, not systematic entry into the administrative reporting system only above the mandate threshold.

As a coverage diagnostic, a formal RD test on the probability of individual reporting yields local-linear means of 0.909 just below 300 ppsm and 0.982 just above — a gap of about seven percentage points on a base above ninety percent — which the bias-corrected estimator does not reject as discontinuous (0.037, $p = 0.563$). This pattern is consistent with the institutional structure described in Section 2: both sides of the cutoff operate inside Act 101's statewide reporting-and-grant system, while mandated municipalities face additional legal obligations. The reporting system is therefore not a voluntary survey in which threshold status mechanically determines whether a municipality appears. Municipality-years without individual tonnage are excluded from the recycling-outcome

[13]Consolidations and mergers under the 1994 Municipal Consolidation Act are rare (Pennsylvania Local Government Commission, 2025, U.S. Census Bureau, 1994): fifteen boundary changes.

[14]Of the 373 municipalities ever appearing in the sample, a fully balanced panel would yield 373×24 muni-years; the observed count is 7,521, leaving 1,431 muni-years outside the population range under the prevailing decennial benchmark (860 below 5,000 and 571 at or above 10,000). These out-of-range municipality-years are not exits from the DEP system: 85 percent (1,211 of 1,431) still file an individual tonnage report and 72 percent (1,024 of 1,431) still operate a curbside program. They are simply outside the analysis window for that decade and re-enter when the next benchmark returns their population to the 5,000–10,000 range.

analysis but retained in the covariate balance table (Table 1); of the 527 excluded, 426 fall below the density threshold and 101 above. I report balance on the full 5,000–10,000 population sample rather than on reporters only, because reporting status is potentially affected by treatment and conditioning on it would test the wrong null; Appendix Table A2 separately confirms balance among the excluded non-reporters. This individual-reporting coverage diagnostic should not be confused with the within-sample activation margin defined in Section 6, which conditions on observed individual municipal tonnage and asks where, within that sample, the discontinuity is generated.

Two further restrictions produce the estimation sample. The first removes 117 municipality-years with implausibly large per capita tonnage as data-quality outliers,[15] leaving a clean panel of 6,877 municipality-years. The second restricts estimation to municipalities near the 300 ppsm threshold, yielding the 3,941 municipality-years (230 unique municipalities) that appear in the main estimation tables. Covariate balance at the density threshold for this estimation sample is reported in Table 1; the validity implications are discussed in Section 4.1.

The 3,941 municipality-years that survive these restrictions are the analysis sample for every estimate that follows: DEP coverage is near-universal within the 5,000–10,000 population band, individual reporting and data-quality flags show no detected discontinuity at the density threshold, and predetermined 1980 covariates are balanced. Effective samples for individual specifications may differ by a small number of singleton-absorbed observations (Table 2). The 300 ppsm assignment rule operates sharply on this sample, which is what the regression discontinuity design in Section 4 requires.

4 Empirical Strategy

Pennsylvania's 300 ppsm density threshold partitions municipalities in the 5,000–10,000 population band into mandated and unmandated arms of the same Act 101 statute. The empirical strategy compares municipalities just above and below this inherited density cutoff within the statewide recycling-finance and reporting environment described in Section 2; the DEP coverage documented in Section 3 makes this comparison observable in municipal tonnage records. This section formalizes the pooled municipality-year RD design, defines treatment assignment by decennial-Census density regime, and estimates proportional effects using PPML. It then evaluates the local validity of the cutoff: predetermined 1980 covariates do not jump at 300 ppsm, first-class township status is locally smooth, and the density distribution shows no evidence of bunching at the threshold.

The design sits inside a literature that uses statutory thresholds in local government to identify institutional effects (Grembi et al., 2016, Eggers et al., 2018), and the standard threat to that literature is bundling: population cutoffs at 5,000 or 10,000 inhabitants typically trigger several institutional changes at once — council size, mayoral salary, fiscal regime, audit requirements —

[15] A municipality-year is flagged if total recycling exceeds 2.0 tons per capita, residential recycling exceeds 1.0 tons per capita, or commercial recycling exceeds 1.5 tons per capita — thresholds approximately 3–4 times the 90th percentile of the cleaned distribution. Inspection and DEP correspondence confirm the dominant mechanism: packaging waste from a distribution center or adjacent manufacturing facility is reported as post-consumer recycling in the municipality's Act 101 submission, inflating the commercial stream. The two groups — flagged and missing-tonnage — are mutually exclusive. The baseline rate of this flag is 1.7 percent for reporting municipalities in the sample and is balanced across the density threshold: a regression discontinuity on the flag indicator at 300 ppsm yields a coefficient of 2.08 percentage points with bias-corrected $p = 0.11$. The point estimate is, if anything, marginally positive — slightly *more* flagging above the cutoff — so removing flagged observations is unlikely to inflate the commercial-stream RD reported in Section 5.

forcing the analyst either to net out the bundle or to settle for the joint effect. Unlike population thresholds, the 300 ppsm density cutoff does not appear to coincide with other statewide municipal obligations in the 5,000–10,000 population band, narrowing the usual bundling concern. The discontinuity provides a clean source of variation in exposure to the recycling mandate.

Treatment assignment. The design is a pooled municipality-year RD with treatment assigned by decennial-Census density regime. Let x_{it} denote population density in municipality i at time t, centered at the statutory threshold:

$$x_{it} = \text{density}_{it} - 300, \qquad D_{it} = \mathbb{1}(x_{it} \geq 0). \tag{1}$$

Density is recalculated at each decennial Census, but only 24 of 373 municipalities in the sample ($\approx 6.4\,\%$) ever change sides across the three Census benchmarks; the identifying variation is therefore the cross-sectional discontinuity at 300 ppsm within each Census regime, not within-municipality changes in treatment status.[16]

Specification. I estimate the proportional jump in recycling at the threshold via the following PPML specification. For municipality i in county c and year t,

$$E[y_{it} \mid \cdot] = \exp\big(\tau D_{it} + \beta_1 x_{it} + \beta_2 (D_{it} \cdot x_{it}) + \gamma' Z_{it} + \lambda_t + \mu_c\big), \tag{2}$$

where y_{it} is DEP-reported recycling tonnage per capita (total, residential, commercial), λ_t are year fixed effects, μ_c are county fixed effects, and Z_{it} is a vector of predetermined municipal controls. Separate slopes β_1 below and $\beta_1 + \beta_2$ above the cutoff implement the local-linear specification of Imbens and Lemieux (2008) and Gelman and Imbens (2019). Observations within $[-h, h]$ receive triangular kernel weights $w_{it} = \max(0,\, 1 - |x_{it}|/h)$. The coefficient τ is the log-point jump at the threshold; $\exp(\tau) - 1$ is the implied proportional change in recycling per capita at the cutoff.

Identifying variation is the cross-sectional density discontinuity at 300 ppsm. Year effects λ_t absorb common statewide shocks — commodity prices, regulatory amendments — and county effects μ_c absorb time-invariant differences in collection infrastructure, MRF access, and county-level enforcement intensity, sharpening precision without altering the source of variation.

Estimator. The benchmark specification estimates proportional effects using PPML. Recycling tonnage is non-negative, right-skewed, and includes a non-trivial mass of zeros — up to 23.4 percent on the control side of the commercial stream. Log-linear OLS is therefore unattractive: using $\log y$ drops zero outcomes, while using $\log(1+y)$ imposes an arbitrary additive constant whose scale dependence changes the estimand (Chen and Roth, 2024). PPML instead models the conditional mean directly with a log link, retains zero outcomes, and is consistent for the mean under multiplicative heteroskedasticity without requiring the outcome to be Poisson distributed (Santos Silva and Tenreyro, 2006, Chen and Roth, 2024). I report treatment effects as $100 \times (\exp(\hat{\tau}) - 1)$, the implied percent change at the cutoff.[17]

[16] A within-switcher specification on the 24 side-changing municipalities is directionally consistent with the pooled estimate but too imprecise to interpret on its own. Results are available upon request.

[17] Table 4 shows that the results are not driven by the PPML functional form. A local-linear RD estimator in levels delivers positive estimates for all three outcomes and statistically significant effects for total and commercial recycling under robust bias-corrected inference. The residential estimate remains positive but is less precise under the bias-corrected local-linear estimate.

Controls and bandwidth. The preferred specification includes Z_{it} for precision; identification rests on continuity at the cutoff rather than on covariate adjustment. The controls are the twelve predetermined 1980 Census characteristics introduced in Section 3, grouped there by the four channels through which Act 101's legislative history flags plausible confounding. The local-linear slopes already absorb smooth variation correlated with density, so consistency does not require balance at the cutoff (Panel B of Table 1 confirms this directly). All twelve covariates are measured in 1980 — eight years before Act 101 and a decade before the panel begins — so none is post-treatment.

Standard errors are clustered at the municipality level.

The common preferred bandwidth is $h = 845$ ppsm, the largest of the three aggregate-specific MSE-optimal symmetric bandwidths for the local-linear estimator in levels (Calonico et al., 2014) (768, 676, and 845 ppsm for total, residential, and commercial recycling). Pooling on the widest of the three MSE-optimal bandwidths keeps cross-aggregate comparisons on a common identifying sample, so differences across total, residential, and commercial reflect differences in response rather than differences in support. Outcome-specific MSE-optimal estimates (768, 676, and 845 ppsm) are reported in Panel B of Table 4 and deliver the same pattern on all three streams. Panel B of Table 4 also shows that the estimates are stable at $h/3$, $h/2$, $2h$, and at the CER-optimal bandwidth.

Identification. The RD estimand is the discontinuous change in municipality-level DEP-recorded recycling at the 300 ppsm threshold for municipalities in the 5,000–10,000 population band. Because D_{it} records statutory mandate status rather than observed implementation, τ is a reduced-form effect of being legally subject to the mandate. Compliance is near-universal, so τ is also close to the effect of mandate compliance; Section 3 verifies that DEP coverage does not jump at the cutoff. Section 5 characterizes the marginal municipalities to which this local estimate applies.

Identification requires continuity of potential outcomes at the cutoff: absent the mandate, recycling outcomes would have varied smoothly with municipal density around 300 ppsm. Density itself is plausibly related to recycling because denser municipalities may face lower average collection costs and higher voluntary curbside adoption (Kinnaman, 2000). That relationship does not invalidate the design so long as it is smooth at the threshold. The local-linear specification therefore allows recycling to vary with density on both sides of the cutoff and attributes only a discrete break at 300 ppsm to the mandate.

The main threats are correspondingly local: another policy or institution could change discontinuously at the same threshold, or municipalities could sort around the cutoff. The statutory history reduces both concerns because the 300 ppsm threshold originated in 1899, nearly nine decades before Act 101, and was not created in response to recycling policy. Section 4.1 evaluates these threats directly by testing for discontinuities in predetermined municipal characteristics and for manipulation of the running variable at the cutoff.

4.1 Validity

Pre-mandate covariate balance. A first implication of the RD identifying assumption is that predetermined municipal characteristics should not change discontinuously at 300 ppsm. I test this implication using the twelve 1980 Census covariates that enter Z_{it} in Equation (2). Because these characteristics are measured eight years before Act 101's passage and a full decade before the recycling panel begins, they cannot be affected by the mandate. The running variable for this

balance test is 1980 density, defined as 1980 NHGIS population divided by 1990 Gazetteer land area, as described in Section 3.

Table 1 reports the local-linear RD discontinuity and robust p-value for each covariate at the baseline bandwidth. None of the twelve predetermined covariates jumps at 300 ppsm: 0 of 12 discontinuities are statistically significant at the 10 percent level. This pattern weighs against secular sorting on pre-existing traits predictive of later recycling capacity.

The raw above-versus-below differences are largest for fiscal capacity: the college share is 0.103 below the cutoff and 0.162 above it, while median household income is \$ 81.8K below and \$ 97.1K above, in 2026 dollars. Locally at the cutoff, however, neither variable exhibits a discontinuity: the college-share RD estimate is 0.010 ($p = 0.723$), and the income estimate is \$ 5.120K ($p = 0.304$). The remaining ten covariates, covering demographic composition, housing structure, and five 1980 industry shares, also show no local break, with every RD p-value above 0.36.

First-class township status. The one institution that shares the 300 ppsm threshold's 1899 statutory origin is first-class township classification (Section 2). It therefore warrants separate attention: if first-class status changed discontinuously at the same cutoff, the RD estimate could conflate the recycling mandate with this older municipal institution. In practice, however, first-class status is not mechanically assigned at 300 ppsm. A township must both satisfy the density requirement and approve the classification by referendum, so eligibility and actual classification diverge. As of Act 101 enactment, only 5.6 percent of treated municipalities and 0.0 percent of control municipalities were first-class townships.

The balance test in Table 1 shows no local discontinuity in actual first-class status at the recycling cutoff: the RD coefficient is 0.059 ($p = 0.310$; Table 1). The main estimates are also unchanged when first-class townships are excluded from the sample (Section 7). These results make it unlikely that first-class township classification is driving the estimated mandate effect.

No manipulation of the running variable. A separate diagnostic for the RD design is whether municipalities bunch discontinuously around 300 ppsm. Strategic manipulation is unlikely in this setting. Density is computed from Census population and municipal land area, and the 300 ppsm threshold originated in 1899, nearly nine decades before Act 101 (Section 2). Any sorting around the threshold would therefore have had to anticipate a much later recycling mandate or operate through slow-moving demographic and boundary changes rather than direct adjustment of the running variable.

Figure 2 examines whether the density distribution changes discontinuously at the cutoff. Panel (a) plots the raw density distribution of Pennsylvania municipalities, and panel (b) implements the Cattaneo et al. (2018) local-polynomial density test on the centred running variable. Although the distribution is visibly asymmetric over the wider window, the relevant diagnostic is whether density changes discontinuously at $x = 0$. It does not: the fitted densities meet smoothly at the cutoff, with no evidence of bunching immediately below or above 300 ppsm.

The Cattaneo et al. (2018) test gives the same conclusion. Using one observation per municipality at its earliest Census-era density value ($N = 373$), the test yields $T = -0.534$ ($p = 0.593$, two-sided). A decade-panel version, with one observation per municipality–Census-regime cell, yields $T = -1.582$ ($p = 0.114$, two-sided; $N = 936$). Thus the continuity result is unchanged whether the running variable is collapsed to one observation per municipality or observed separately by Census regime.

5 Results

This section shows that DEP-recorded municipal recycling tonnage per capita increases discontinuously at the 300 ppsm threshold. In the preferred specification, the RD estimate implies that threshold-adjacent municipalities subject to the mandate have 99.3 percent more total recycling per resident per year than municipalities just below the cutoff, equivalent to roughly 199 additional pounds per capita at the below-threshold mean. The estimated increase appears in both disaggregated outcomes: residential recycling is 53.1 percent higher and commercial recycling is 147.2 percent higher.

Graphical preview. The discontinuity is visible in the raw DEP data. Figure 3 plots binned averages of recycling tonnage per capita against centered density, with separate linear fits on either side of the cutoff. Municipalities just above 300 ppsm report substantially more recycling per capita than municipalities just below it, and the fitted trends do not suggest that the difference is driven by a gradual change in density away from the cutoff. The estimated break is large relative to within-side variation in conditional means. As a falsification exercise, Figure 5 reports RD estimates at artificial cutoffs in the 100–200 and 400–700 ppsm ranges, each using a window that excludes the true cutoff; none shows a break comparable to the estimate at 300 ppsm.

Benchmark total-recycling estimate. The estimated discontinuity is stable across the specification ladder. Panel A of Table 2 reports a 68.6 percent increase in total recycling per capita in the no-controls specification (column 1), 67.0 percent after adding year fixed effects (column 2), 89.1 percent after adding county fixed effects (column 3), and 99.3 percent after adding the full set of predetermined 1980 Census controls (column 4). All four 95% confidence intervals exclude zero. Column 4 is the preferred specification.

The 1980 controls are predetermined: they are measured at least eight years before Act 101 and at least twenty-one years before the first recycling observation in the panel. Their inclusion adjusts for cross-sectional differences in pre-mandate economic structure without conditioning on post-treatment outcomes. Holding the year and county fixed effects at the column-(4) structure and varying only the control set, joint Wald tests cannot reject equality of the treatment coefficient across the five nested control sets for any of the three outcomes ($p > 0.55$ for total, residential, and commercial recycling; Appendix Table B1). A mean-based translation puts the preferred estimate at roughly 199 additional pounds per resident per year among municipalities near the cutoff, evaluated at the below-threshold mean of 200.0 pounds per capita per year. This level comparison is not an artifact of dropping zero-tonnage observations. Zero-valued tonnage observations are concentrated on the control side— 4.2 percent of control municipality-years versus 0.0 percent of treated municipality-years—and PPML retains these observations through the log-linear conditional mean, whereas a log-linear OLS specification would discard them. Because the panel begins in 2001, thirteen years after Act 101's 1988 enactment, the estimate should be interpreted as a sustained difference associated with mandate exposure rather than an immediate impact effect.

The preferred estimate is local to municipalities near the statutory cutoff, rather than an average effect for all Pennsylvania municipalities or all U.S. communities. To characterize that local population, Table B2 compares the Abadie kappa-weighted RD sample with the U.S. suburban-transition band excluding Pennsylvania. The weighted RD sample closely matches the comparison band on college share (30.6 vs. 30.6 percent; difference 0.0 pp, SE 1.1), but differs on the other observables: it is whiter (89.4 vs. 81.3 percent; difference +8.0 pp, SE 0.7), has lower unemployment (3.9 vs. 4.8 percent; difference -0.9 pp, SE 0.2), and has substantially higher median household

income ($101,052 vs. $80,923; difference $20,129, SE $2,620). The local RD estimate should therefore be interpreted as applying to a whiter, lower-unemployment, and more affluent subset of the suburban-transition band around the 300 ppsm cutoff.

Residential and commercial recycling. Both outcomes increase at the cutoff. Panels B and C of Table 2 show that residential recycling per capita is 53.1 percent higher among threshold-adjacent treated municipalities in the preferred specification (Panel B, column 4; control mean 92.3 lbs), while commercial recycling per capita is 147.2 percent higher (Panel C, column 4; control mean 107.7 lbs). Both estimates are positive and statistically significant across the full specification ladder: no fixed effects, year fixed effects, county fixed effects, and the full-control benchmark.[18] In mean-based level translations, the residential estimate corresponds to approximately 49 additional pounds per capita per year and the commercial estimate to approximately 158 additional pounds, summing to roughly 207 pounds.[19]

A stacked PPML specification pools residential and commercial municipality-year observations, includes a source indicator and a treatment × source interaction, allows separate running-variable slopes and year fixed effects by source, and estimates both responses at the common bandwidth. The interaction test cannot reject equality of the two responses ($p = 0.487$; bootstrap $p = 0.595$). Although the point estimates differ— 53.1 percent for residential recycling and 147.2 percent for commercial recycling—the data do not distinguish the responses at conventional levels.

Persistence and comparison to prior work. The discontinuity is persistent over the sample period. Figure 1 plots average total recycling per capita for municipalities above and below 300 ppsm, from 2001 to 2024. The gap is already present at the start of the sample period—thirteen years after enactment—and remains visible over the next two decades. Below-threshold municipalities do not visibly catch up, and year-to-year movements affect both groups in broadly similar ways, consistent with common shocks to recycling markets and statewide reporting practices.

The formal decade-specific estimates corroborate this average pattern. Table 3 re-estimates the benchmark specification separately for 2001–2009, 2010–2019, and 2020–2024, corresponding to the 2000, 2010, and 2020 Census decades. Total recycling is 122.1, 132.8, and 85.2 percent higher across the three decades, bracketing the pooled benchmark of 99.3 percent ($p = 0.810$ for equality across decades). Residential recycling is 78.8, 68.2, and 43.7 percent higher ($p = 0.711$), and commercial recycling is 216.2, 205.3, and 138.0 percent higher ($p = 0.451$). Every decade-specific estimate is positive and statistically significant, and the Wald tests do not reject equality across decades for any outcome. Because these periods span major shifts in recyclables markets—including the rise of Chinese demand for secondary materials through 2009 and the disruption following China's 2018 National Sword restrictions—the stability in Table 3 is consistent with a persistent mandate discontinuity across different commodity-market environments.

The three estimates admit different comparisons with the literature. The commercial response speaks to nonresidential recycling tonnage, a margin with little direct quasi-experimental evidence. The residential response can be compared to the household-elasticity literature and, more directly, to household-panel estimates of state recycling-law exposure.

[18] Zeros are more prevalent in the decomposed series than in the total— 8.5 vs. 0.8 percent of municipality-years for residential recycling and 23.4 vs. 2.5 percent for commercial recycling—which reinforces the use of PPML rather than a log-linear estimator that would discard zero observations.

[19] The residential estimate does not separate extensive-margin effects, such as new household participation, from intensive-margin effects, such as higher volumes among existing participants. Direct participation-rate data would be required to isolate those household-level channels.

Starting with the commercial estimate, the closest analogues measure adjacent outcomes and rely on broader policy-regime contrasts. Ogieriakhi and Wang (2024) compare California after AB 341 with Washington, estimating the effect of a commercial recycling mandate on landfilled waste rather than recycling tonnage. Anglou et al. (2024) study state commercial food-waste bans using landfill outcomes, and Ishimura (2022) studies Japan's Containers and Packaging Recycling Law under a national producer-financed regime rather than a separately observed commercial recycling outcome. These designs are informative about nonresidential waste and packaging policy, but they recover policy changes embedded in cross-state or national administrative, market, and enforcement environments. By contrast, the RD design here holds the Pennsylvania policy regime fixed and isolates assignment to the municipal recycling obligation at a statutory cutoff. To my knowledge, this is the first quasi-experimental estimate of a recycling mandate's effect on commercial recycling tonnage in a long municipal panel.

The broader causal literature is more developed on the residential side, where it generally identifies a different object: household responses to marginal price, access, or program-feature variation, holding the municipal recycling obligation fixed (Fullerton and Kinnaman, 1996, Kinnaman, 2006, Jenkins et al., 2003, Allers and Hoeben, 2010). A related policy-design literature compares municipal solid waste instruments such as deposit-refund systems, advance disposal fees, and recycling subsidies (Palmer et al., 1997). The mandate studied here varies whether a municipal program is legally required, which materials it must cover, and whether tonnage reporting is part of a compliance obligation—a municipal-level object, of which the household elasticity is one component.

The closest household-policy comparison is Viscusi et al. (2023), who first-difference 232,309 pairs of same-household survey observations and use changes in the household's recycling environment to explain changes in self-reported recycling of paper, cans, glass, and plastic. Their outcome is a material count, not tonnage or program provision. Their effective-recycling-law measure pools mandatory recycling statutes with opportunity-to-recycle and recycling-plan statutes; the coefficient comes from households that move into a state with such a law and is 0.18 materials, about 6% relative to the 2.79-material mean. The RD estimate here is larger because it identifies a different margin: within a fixed Pennsylvania state-law regime, assignment of a municipality to an obligation that changes program provision, material coverage, and compliance reporting. Viscusi et al. (2023) therefore estimate a mover-based household sorting response to a bundled state-law environment, not the effect of assigning a local government to a statutory collection obligation.

The magnitude raises the question of *how* the additional tonnage is generated. Section 6 decomposes the aggregate response along three margins of adjustment—activation, composition, and intensification—and along the statutory material partition, separating non-covered tonnage, non-leaf covered materials, and leaf waste (covered under Act 101's yard-waste requirement but institutionally distinct).

6 Decomposing the Mandate Effect

The estimates in Section 5 show that mandate exposure sharply increases DEP-recorded municipal recycling at the 300 ppsm cutoff. This section asks where that increase appears within the administrative records. I decompose the cutoff jump into three nested margins: *activation*, whether a municipality-year records any positive tonnage; *composition*, whether material groups appear with positive tonnage; and *intensification*, tonnage within active material groups. The activation margin is distinct from the reporting-coverage diagnostic in Section 3, which shows high individual

municipal reporting and no detected discontinuity at the cutoff.

The decomposition disciplines interpretation. A response concentrated in activation or material-group appearance would point toward administrative entry or broader recorded scope. A response concentrated in intensification, especially within materials covered by Act 101, is harder to explain as reporting entry or category appearance alone. This distinction is reinforced by the institutional setting: as shown in Section 2, the main per-ton grant incentive does not begin at 300 ppsm. Municipalities on both sides of the cutoff can report tonnage to DEP and receive performance grants, while the cutoff changes mandatory compliance with the Act 101 bundle. Covered-material intensification therefore captures the part of the response least exposed to changes in reporting coverage or category appearance.

Decomposition and identification. Appendix C derives the decomposition; the main text uses the resulting accounting identity. For each outcome—total, residential, and commercial recycling—let y_{it} denote per capita recorded tonnage. I partition materials into three groups, $\mathcal{K}_3 = \{\ell, m, n\}$: ℓ is leaf waste, covered through Act 101's yard-waste requirement; m is the remaining non-leaf material covered by the relevant Act 101 obligation; and n is non-covered tonnage. Let $a_{it} \equiv \mathbb{1}\{y_{it} > 0\}$ denote activation, and let $g_{kit} \equiv \mathbb{1}\{y_{kit} > 0\}$ indicate whether material group k appears among municipality-years with positive tonnage. Then

$$y_{it} = a_{it} \sum_{k \in \mathcal{K}_3} g_{kit}\, y_{kit},$$

where y_{kit} is per capita tonnage within group k when that group appears.

Because y_{it} is a sum across material groups, a log discontinuity does not decompose exactly except in special cases. I therefore use a first-order log decomposition around control-side material shares. The approximation applies the standard first-order expansion $\Delta \log(\sum_k x_k) \simeq \sum_k \omega_k \Delta \log x_k$, where the weights are evaluated at the control-side shares and the residual below is the difference between the PPML reduced-form effect and this share-weighted sum. Let ω_k be group k's share of per capita recyclables among activated control municipality-years within the RD bandwidth, pooling all sample years. The PPML reduced-form effect can be approximated by

$$\tau = \underbrace{\tau_a}_{\text{activation}} + \underbrace{\omega_\ell \tau_\ell^g + \omega_m \tau_m^g + \omega_n \tau_n^g}_{\text{composition}} + \underbrace{\omega_\ell \tau_\ell^y + \omega_m \tau_m^y + \omega_n \tau_n^y}_{\text{intensification}} + \varepsilon. \tag{3}$$

Each component is in log points. The term τ_a is the cutoff jump in the fitted probability of positive recorded tonnage. Conditional on activation, τ_k^g is the cutoff jump in the fitted probability that material group k appears, and τ_k^y is the cutoff jump in fitted per capita tonnage within group k when it appears. The residual ε is the approximation error from applying the share-weighted log decomposition to a sum of material groups.[20]

The covered group m is defined separately by outcome. For residential tonnage, it contains the statutory curbside material families: paper, glass, metal cans, and eligible plastics. Single-stream and commingled aggregates are also assigned to m because the DEP aggregate categories primarily bundle these curbside materials rather than separately identifying their statutory components. These categories account for roughly two-thirds of control-side residential tonnage. For commercial tonnage, m contains the statutory commercial set: high-grade office paper, corrugated paper,

[20]The estimated approximation error is small: $\hat{\varepsilon}_{\text{tot}} = 0.002$, $\hat{\varepsilon}_{\text{res}} = -0.015$, and $\hat{\varepsilon}_{\text{com}} = -0.043$ log-points. Table C1 reports the residuals, and Appendix C.2 describes their construction.

and aluminum cans. The non-covered group n contains all remaining categories, including non-bottle plastics, non-container glass, batteries, electronics, household hazardous waste, textiles, and organics.

The composition and intensification terms condition on selected samples, so I report Lee bounds under a one-sided selection assumption within each partition cell. The assumption allows the mandate to induce positive tonnage or positive material-group appearance, but rules out mandate-induced exit from the relevant selected sample. This restriction is most natural for activation and covered materials, where Act 101 creates affirmative obligations, and more demanding for non-covered materials, where the statute neither requires nor forbids reporting. I therefore treat non-covered estimates more cautiously. Appendix C.3 states the assumption, reports the bounds, and shows sensitivity to limited monotonicity violations.

Aggregate composition and intensification use weighted rectangular Lee bounds, formed by summing the component lower endpoints and upper endpoints separately. These bounds are conservative because they allow the worst-case selection scenario to occur simultaneously in the covered and non-covered groups. I use these rectangular bounds as the main object because they bound the same additive aggregate in Equation (3) using only the cell-wise one-sided selection restrictions. Sharper joint bounds would either target a joint-active subpopulation or require additional restrictions on how selection is coupled across material-group cells.

Full sample and curbside subsample. I estimate the decomposition using the column-(3) specification from Table 2: separate slopes, a triangular kernel, and year and county fixed effects. This is the specification compatible with Lee trimming on the level activation probability; Appendix C.5 show that the point-identified decomposition is unchanged when the full continuous-control specification is used.

I first use the full RD sample, then repeat the exercise among municipalities with curbside collection recorded in 2001, the first year of the outcome panel. The curbside restriction is an interpretive check, not a separate causal design, because 2001 curbside status is measured after Act 101 took effect. It asks whether the full-sample pattern is mechanically driven by municipalities without any recorded curbside program at the start of the panel. Baseline curbside status is smooth at the cutoff, as shown in Appendix C.6, so the restriction does not itself introduce a discontinuous change in measured sample composition.[21]

Decomposition results. Figure 4 provides the section's main evidence on the margins through which the mandate raises recorded recycling. Each row reports one outcome—total, residential, and commercial recycling. Within each row, the left panel uses the full RD sample and the right panel repeats the same decomposition among municipalities with curbside collection recorded in 2001. The bars report point estimates and split the total cutoff effect into activation, material-group composition, and material-group intensification. The uncertainty intervals differ by margin. For activation, the intervals are conventional 95 percent confidence intervals for the RD coefficient on positive recorded tonnage; because the activation estimates are very small, these intervals are often visually indistinguishable from the point estimates. For composition and intensification, the vertical intervals are Lee bounds for selection into positive recorded tonnage or positive material-group tonnage, not sampling confidence intervals. Tables C4 and C6 report all estimates and

[21] This smoothness should not be read as evidence that the mandate was irrelevant to curbside recycling. The 2001 curbside measure is far downstream of Act 101 and the post-enactment diffusion of curbside service; Appendix C.6 summarizes the DEP counts.

bounds numerically.

In the full sample, activation explains essentially none of the discontinuity. The activation components are small and not statistically significant at the 5 percent level: 0.006 log points for total recycling, 0.018 for residential, and 0.040 for commercial, compared with reduced-form effects of 0.637, 0.436, and 0.795. The composition margins show that the mandate broadens recorded material scope. Composition is positive across material groups in the total and residential decompositions, and positive for covered and non-covered materials in the commercial decomposition; commercial leaf composition is economically negligible and not sign-identified under Lee bounds. These recorded-scope margins matter, especially for residential leaf waste, but they are not the main source of the total or commercial effects.

The largest total and commercial contributions come from intensification. For total recycling, intensification accounts for 0.52 log points of the 0.64 log-point discontinuity, including 0.31 log points from non-leaf Act 101-covered materials; the Lee interval for that covered component is [0.19, 0.33]. Commercial recycling gives the clearest version of the pattern: 0.70 log points of the 0.80 log-point effect come from intensification, including 0.42 from non-leaf covered materials, with Lee interval [0.10, 0.48]. Residential recycling is more mixed: leaf composition contributes materially, and non-leaf covered intensification is smaller, 0.10 log points, though positive under Lee trimming, [0.01, 0.13]. This difference is not surprising: control-side activity is denser for covered materials, while many non-covered categories have sparse baseline reporting, making their Lee intervals more sensitive to selection.

The curbside subsample sharpens the interpretation. Total and commercial effects remain large, at 0.56 and 0.73 log points, with sizable non-leaf covered intensification components of 0.27 and 0.42. Residential recycling changes more sharply: the reduced-form effect falls to 0.33 log points, and non-leaf covered intensification is only 0.01, with Lee bounds [-0.06, 0.04]. The residential full-sample effect should therefore not be read as a large additional household sorting response among municipalities that already had curbside service. It is better interpreted as the long-run effect of the broader residential compliance bundle: curbside provision, leaf-waste collection, material scope, public education, compliance, and reporting formalization.

Relation to household-facing evidence. This pattern clarifies how the residential estimates relate to household-facing recycling evidence. Prior work shows that households respond when policy changes the cost, convenience, or salience of recycling, including through curbside access, unit pricing, deposit-refund systems, and sorting rules (Jenkins et al., 2003, Best and Kneip, 2019, Fullerton and Kinnaman, 1996, Carattini et al., 2018, Viscusi et al., 2012, 2023, Akbulut-Yuksel and Boulatoff, 2021). The comparison here is therefore one of scale and margin, not a common elasticity. The full-sample residential effect is nontrivial because Act 101 changes the recycling regime households face. But the curbside-subsample estimate disciplines that interpretation: among municipalities with curbside service already recorded in 2001, the additional residential non-leaf covered-material response is small. Consistent with evidence that curbside expansion can have limited marginal effects when recycling alternatives or baseline programs already exist (Kinnaman, 2006, Beatty et al., 2007), Act 101's long-run household-side effect appears to operate mainly through the program bundle—service availability, material scope, leaf-waste requirements, compliance, and formalization—rather than through a large marginal increase in sorting intensity once curbside collection already exists.

Commercial response. The commercial panels identify a margin largely outside the household curbside literature. Existing causal evidence on recycling policy is concentrated on households, while commercial recycling mandates and related waste regulations are typically evaluated using state-level policy variation or landfill-disposal outcomes (Ogieriakhi and Wang, 2024, Anglou et al., 2024). Here, the DEP data separately identify commercial tonnage, and the RD design holds fixed the statewide legal, grant-eligibility, and reporting environment.

The key institutional fact is that Act 101 does not require municipalities to collect commercial recycling directly. Commercial generators may continue using private haulers; the mandate instead requires commercial, municipal, and institutional generators to separate the statutory commercial set and requires municipalities to document that tonnage in the Act 101 filing. The decomposition maps onto this statutory margin: the commercial effect is large in the full sample and remains large in the 2001-curbside subsample; activation is small; composition contributes, but the largest contribution is covered-material intensification. Mandated municipalities report more tonnage in statutory commercial materials already reported with positive tonnage on both sides of the cutoff.

I interpret the commercial estimate as a durable increase in covered commercial recycling recorded through the municipal compliance system. It may reflect additional source separation by newly covered commercial generators, more complete documentation of private-hauler recycling that was already occurring, or both. Because the records cannot distinguish those channels, the narrower interpretation is covered commercial compliance intensification within the municipal reporting system.

Summary. The decomposition yields three main conclusions. First, activation is small, so the aggregate discontinuity is not driven by municipalities moving from zero to positive recorded recycling. Second, the residential response is substantial in the full sample but attenuates in the 2001-curbside subsample, pointing to the broader residential compliance bundle rather than a large additional household sorting response conditional on curbside service. Third, the commercial response remains large in the curbside subsample and is centered on covered-material intensification within the municipal compliance system.

7 Robustness

Two standard RD validity checks are reported in Section 4: covariate balance on pre-mandate Census characteristics in Table 1, and the Cattaneo et al. (2018) density test in Figure 2. This section examines whether the main estimate is sensitive to specification, sample, estimator, and inference choices.

Table 4 reports the robustness exercises. Panels A–F vary the estimating sample or point-estimate specification; Panels G–H vary the inference procedure. Across the PPML specifications, the total recycling estimate remains positive and statistically significant at $p < 0.05$, ranging from 51.4 percent at one-third of the MSE-optimal bandwidth to 122.9 percent under the widest donut restriction. Residential and commercial recycling display the same sign pattern, although the residential estimate is less precise in the narrowest-window and bias-corrected local-linear specifications.

Bandwidth variation (Panel B). Panel B first replaces the common baseline bandwidth, $h = 845$, with outcome-specific MSE-optimal bandwidths: 768, 676, and 845 ppsm for total, residential, and commercial recycling, respectively. The resulting estimates remain positive and

significant at $p < 0.01$ for all three outcomes, with implied effects of 93.1, 44.7, and 147.2 percent. Thus, the baseline result does not depend on imposing a common bandwidth across outcomes.

The panel then varies the common bandwidth to one-third, one-half, and twice the MSE-optimal value, and also reports the CER-optimal bandwidth of Calonico et al. (2014). Total recycling effects range from 51.4 percent at $h/3$ to 115.4 percent at $2h$. The CER-optimal estimate is 81.8 percent, below the MSE-optimal baseline. This comparison changes both the bias–coverage tradeoff and the realized common-support sample used to decompose total recycling into residential and commercial components. Across the bandwidth grid, the total and commercial estimates remain positive and significant. Residential recycling is also positive throughout, but the estimate at $h/3$ is smaller and not conventionally significant (15.9 percent, $p > 0.10$), consistent with the loss of precision in the narrowest window.

Kernel robustness (Panel C). Panel C replaces the triangular kernel with a rectangular kernel. The total recycling estimate increases from 99.3 percent in the baseline to 119.5 percent, with little change in precision. Residential and commercial estimates also remain positive and significant at $p < 0.01$.

Local-linear RD robustness check (Panel D). Panel D estimates a local-linear RD specification following Calonico et al. (2014), using the baseline common bandwidth and triangular kernel, with year and county dummies included as additive covariates. Because the local-linear coefficient is measured in levels, the table rescales each estimate as a percent of the control-side mean within the bandwidth. The conventional estimate, $\hat{\tau}_{cl}$, implies effects of 97.2 percent for total recycling, 58.0 percent for residential recycling, and 130.9 percent for commercial recycling. The corresponding p-values are below 0.01 for total recycling and below 0.05 for the two components.

The robust bias-corrected estimate, $\hat{\tau}_{bc}$, is smaller and less precise, as expected when bias correction is paired with the robust standard error (Calonico et al., 2018). It implies effects of 78.9 percent for total recycling and 129.8 percent for commercial recycling, both significant at $p < 0.05$. The residential estimate falls to 19.5 percent and is not statistically significant. Taken together, these estimates indicate that the discontinuity is not an artifact of the Poisson link. The residential bias-corrected estimate is the noisiest of the three, consistent with the larger robust standard error attached to the smoother residential series. PPML remains the preferred baseline because it accommodates zero commercial recycling observations without imposing an arbitrary log shift (Chen and Roth, 2024).

Donut-hole restrictions (Panel E). Panel E excludes municipality-years within 10, 20, and 30 persons per square mile of the threshold. The total recycling estimate is 89.1 percent with the ± 10 donut, 91.5 percent with the ± 20 donut, and 122.9 percent with the ± 30 donut. The estimates remain statistically significant throughout. The effect therefore does not appear to be driven by municipalities lying immediately adjacent to the 300 ppsm cutoff.

Municipal classification (Panel F). A separate concern is that the density threshold for Act 101 also coincides with the eligibility threshold for Pennsylvania second-class townships to become first-class townships. First-class townships have broader statutory authority, including expanded planning powers and fewer restrictions on service provision, which could independently affect recycling program development. Panel F therefore excludes the ninety-one municipalities

that were first-class townships when Act 101 was enacted, while retaining the baseline bandwidth.[22] The implied effect on total recycling is 104.6 percent, compared with 99.3 percent in the baseline. Residential and commercial estimates are 55.1 percent and 160.1 percent, respectively, and all three estimates are significant at $p < 0.01$. The baseline result is therefore not driven by municipalities that already held first-class status at enactment.

County-clustered inference (Panel G). The baseline specification includes county and year fixed effects, so comparisons are made within counties and common years. This absorbs time-invariant county differences, but it does not by itself determine the appropriate level of clustering. Panel G therefore re-clusters standard errors at the county level, with 42 clusters and an average of 5.23 municipalities per county within the baseline bandwidth. Point estimates are unchanged by construction. The total recycling standard error increases from 0.195 to 0.237 (+22 percent), and the commercial standard error increases from 0.305 to 0.363 (+19 percent). The commercial estimate falls from $p < 0.01$ to $p < 0.05$, while all three outcomes remain statistically distinguishable from zero.

Conley spatial standard errors (Panel H). Panel H reports Conley (1999) spatial HAC standard errors computed from a WLS-FWL linearization with a Bartlett kernel, using 25 km and 50 km cutoff distances. The spatial standard errors are smaller than the municipality-clustered baseline: for example, at 25 km they are 0.166 for total recycling, 0.091 for residential recycling, and 0.267 for commercial recycling, compared with baseline standard errors of 0.195, 0.137, and 0.305. The estimates remain significant at $p < 0.01$ for all three outcomes. The smaller Conley standard errors are consistent with limited residual spatial dependence beyond the municipality once county fixed effects absorb broader spatial structure. They do not suggest that baseline inference is inflated.

Placebo cutoffs. The 300 ppsm cutoff is the only density threshold in the 5,000–10,000 population band at which Act 101 changes municipal obligations. Figure 5 therefore re-estimates the benchmark PPML specification at artificial cutoffs every 25 ppsm in the 100–200 and 400–700 ppsm ranges, excluding a ±100 ppsm neighborhood around the true cutoff. Table B3 reports the corresponding coefficients, bandwidths, and local sample sizes.

The placebo design avoids mechanically comparing observations separated by the true policy threshold. For artificial cutoffs below 300 ppsm, estimation uses only control-side observations; for artificial cutoffs above 300 ppsm, it uses only treated-side observations (Cattaneo et al., 2020). The placebo bandwidth is

$$h_c = \min\{h_c^{\text{MSE}}, |c - c^*| - 1\},$$

with a 100 ppsm minimum distance from the true cutoff. Cutoffs at 200 and 400 ppsm are infeasible under this rule and are dropped, leaving sixteen artificial cutoffs plus the true threshold.

The true cutoff is the largest estimate in absolute rank for total, residential, and commercial recycling. On the 17-cutoff grid, the smallest possible rank statistic is $1/17 \approx 0.059$, which the

[22] Two additional municipalities reclassified to first class during the panel period (Lower Macungie in 2008, Pocono in 2014), but neither is in the RD sample at the time of reclassification: Lower Macungie sits above the 10,000-population ceiling in every year, and Pocono exits the 5,000–10,000 band in 2010, four years before its reclassification. Panel F therefore neither excludes nor relies on these two municipalities for the post-reclassification period. Coding follows township-by-township verification of reclassification years from the Local Geohistory Project.

300 ppsm estimate attains for all three outcomes. Among the 48 placebo coefficients, three exceed $|t| = 1.96$, close to the 2.4 exceedances expected under 48 independent 5% tests, and none survives Benjamini–Hochberg adjustment at $q = 0.05$.

These ranks are diagnostic rather than formal randomization-inference p-values, because the placebo estimates are not exchangeable with the true-cutoff estimate: placebo windows are side-restricted and often use smaller realized bandwidths. The conclusion is unchanged when the grid is restricted to artificial cutoffs with at least 30 municipalities on each side. That restriction leaves ten feasible placebos, and the true cutoff remains rank 1 of 11 for all three outcomes ($p = 0.091$). The evidence is therefore concentrated at the statutory threshold rather than at nearby non-policy densities.

8 Conclusion

Among municipalities near Pennsylvania's 300 persons-per-square-mile cutoff, Act 101 caused large increases in recycling recorded in the DEP municipal system. Using a threshold fixed by historical population density and inherited from law nearly a century older than the recycling statute, I estimate that mandate status raises total recycling per capita by 99.3 percent at the cutoff. At the below-cutoff mean of 200.0 pounds per resident per year, this corresponds to about 199 additional pounds per capita. The corresponding effects are 53.1 percent for residential recycling and 147.2 percent for commercial recycling.

The estimates capture a long-run response. The treated-control gap is already present when the administrative panel begins in 2001, more than a decade after enactment, and remains visible through 2024. Estimates are positive across the 2000, 2010, and 2020 Census regimes, despite large changes in recycling markets, including the post-2018 National Sword disruption. These patterns indicate that Act 101 changed the durable organization of local recycling systems.

A simple reporting-entry interpretation is unlikely to explain the discontinuity. Municipalities on both sides of the cutoff can report eligible tonnage and receive Act 101 performance grants, so the threshold does not create a reporting incentive only for mandated municipalities. Consistent with that institutional setting, the decomposition shows that activation is too small to explain the effect. The main total and commercial responses instead appear within material groups already active on both sides of the cutoff.

The residential and commercial margins point to different mechanisms. For residential recycling, the full-sample effect includes broader recorded scope, especially leaf waste, but the non-leaf covered-material response is small among municipalities with curbside service already recorded in 2001. The residential effect is therefore best read as the long-run effect of the broader compliance bundle—curbside provision, leaf-waste collection, material scope, public education, and reporting formalization—rather than as a large additional household sorting response conditional on curbside service.

For commercial recycling, the strongest signal is higher tonnage within covered commercial materials already reported on both sides of the cutoff. That pattern is consistent with covered commercial compliance intensification within the municipal reporting system. Both channels—additional source separation by commercial generators and more complete documentation of private-hauler flows—are part of the institutional margin Act 101 targets: making commercial recycling obligations observable and enforceable through municipal reports.

The results are local to municipalities near the 300 ppsm cutoff in the 5,000–10,000 population band. They do not speak directly to dense urban centers or sparse rural areas, and the magnitudes

need not extrapolate to jurisdictions with different reporting, grant, or enforcement incentives. What is more likely to travel is the mechanism: recycling mandates can operate through the administrative systems that define covered materials, assign generator obligations, collect records, and make recycling activity visible.

For current packaging-policy debates, including EPR, the implication is that financing rules alone are unlikely to determine performance. Mandates appear most consequential when they create a compliance system: clear covered-material obligations, residential collection coverage, generator duties, material-level reporting, and records that allow municipalities and regulators to see whether obligated flows are actually being captured. This matters most where state laws create producer-side reporting through producer responsibility organizations, while verification of downstream recovery, especially commercial flows handled by private haulers, depends on the municipal and state reporting infrastructure already in place. The Pennsylvania results show that this infrastructure is not automatic: it emerges when policy names covered materials, assigns generator obligations, and requires municipal records that make those flows visible. Without a matching back-end measurement system, producer-financed recycling targets are difficult to verify, particularly on the commercial margin where private-hauler flows are otherwise opaque. The broader lesson is that recycling responsibility becomes enforceable only when policy can follow the material.

References

Akbulut-Yuksel, M. and C. Boulatoff (2021). The effects of a green nudge on municipal solid waste: Evidence from a clear bag policy. *Journal of Environmental Economics and Management 106*, 102404.

Allers, M. A. and C. Hoeben (2010). Effects of unit-based garbage pricing: A differences-in-differences approach. *Environmental and Resource Economics 45*, 405–428.

Anglou, F. Z., R. E. Sanders, and I. Stamatopoulos (2024). Of the first five US states with food waste bans, Massachusetts alone has reduced landfill waste. *Science 385*(6714), 1236–1240.

Beatty, T. K. M., P. Berck, and J. P. Shimshack (2007). Curbside recycling in the presence of alternatives. *Economic Inquiry 45*(4), 739–755.

Best, H. and T. Kneip (2019). Assessing the causal effect of curbside collection on recycling behavior in a non-randomized experiment with self-reported outcome. *Environmental and Resource Economics 72*(4), 1203–1223.

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association 113*(522), 767–779.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Carattini, S., A. Baranzini, and R. Lalive (2018). Is taxing waste a waste of time? Evidence from a Supreme Court decision. *Ecological Economics 148*, 131–151.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge: Cambridge University Press.

Cattaneo, M. D., M. Jansson, and X. Ma (2018). Manipulation testing based on density discontinuity. *The Stata Journal 18*(1), 234–261.

Chen, J. and J. Roth (2024). Logs with zeros? Some problems and solutions. *Quarterly Journal of Economics 139*(2), 891–936.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–45.

Eggers, A. C., R. Freier, V. Grembi, and T. Nannicini (2018). Regression discontinuity designs based on population thresholds: Pitfalls and solutions. *American Journal of Political Science 62*(1), 210–229.

Fullerton, D. and T. C. Kinnaman (1996). Household responses to pricing garbage by the bag. *American Economic Review 86*(4), 971–984.

Gelman, A. and G. Imbens (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics 37*(3), 447–456.

Greenstone, M. and R. Hanna (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review 104*(10), 3038–3072.

Grembi, V., T. Nannicini, and U. Troiano (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics 8*(3), 1–30.

Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*(2), 615–635.

Ishimura, Y. (2022). The effects of the containers and packaging recycling law on the domestic recycling of plastic waste: Evidence from Japan. *Ecological Economics 201*, 107535.

Jenkins, R. R., S. A. Martinez, K. Palmer, and M. J. Podolsky (2003). The determinants of household recycling: A material-specific analysis of recycling program features and unit pricing. *Journal of Environmental Economics and Management 45*(2), 294–318.

Joltreau, E. (2022). Extended producer responsibility, packaging waste reduction and eco-design. *Environmental and Resource Economics 83*(3), 527–578.

Kinnaman, T. C. (2000). Explaining the growth in municipal recycling programs. *Public Works Management & Policy 5*(1), 37–51.

Kinnaman, T. C. (2006). Examining the justification for residential recycling. *Journal of Economic Perspectives 20*(4), 219–232.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies 76*(3), 1071–1102.

National Academies of Sciences, Engineering, and Medicine (2025). *Municipal Solid Waste Recycling in the United States: Analysis of Current and Alternative Approaches*. Washington, DC: The National Academies Press.

National Solid Wastes Management Association (1991). Recycling in the States: 1990 Review. Technical report, National Solid Wastes Management Association, Washington, DC.

OECD (2024). Extended producer responsibility: Basic facts and key principles. OECD Environment Policy Papers 41, OECD Publishing, Paris.

Ogieriakhi, M. and H. Wang (2024). Do mandatory environmental policies really work? A case study of California's mandatory commercial recycling law. *Economic Analysis and Policy 81*, 915–930.

Palmer, K., H. Sigman, and M. Walls (1997). The cost of reducing municipal solid waste. *Journal of Environmental Economics and Management 33*(2), 128–150.

Pennsylvania Department of Environmental Protection (1998). Recycle pennsylvania: Annual report, fiscal year 1997–98. Annual report, act 101 of 1988, Pennsylvania Department of Environmental Protection, Harrisburg, PA.

Pennsylvania Department of Environmental Protection (2001). Pennsylvania recycling program: Annual report, fiscal year 2000–2001. Annual report, act 101 of 1988, Pennsylvania Department of Environmental Protection, Harrisburg, PA.

Pennsylvania Department of Environmental Resources (1991a). Annual report on the pennsylvania recycling program, fiscal year 1990–91. Annual report, act 101 of 1988, Pennsylvania Department of Environmental Resources, Harrisburg, PA.

Pennsylvania Department of Environmental Resources (1991b, April). Status of recycling: Report to the general assembly. Report to the general assembly, Pennsylvania Department of Environmental Resources, Harrisburg, PA. Act 101 of 1988 status report.

Pennsylvania Local Government Commission (2025). Pennsylvania legislator's municipal deskbook. Seventh edition, Harrisburg, PA.

Santos Silva, J. M. C. and S. Tenreyro (2006). The log of gravity. *Review of Economics and Statistics 88*(4), 641–658.

Seif, J. M. (1997, March). Testimony before the Senate Committee on Environment and Public Works. Secretary, Pennsylvania Department of Environmental Protection. 105th Congress.

Shapiro, J. S. and R. Walker (2018). Why is pollution from U.S. manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review 108*(12), 3814–3854.

Sigman, H. and R. Strow (2024). China's waste import ban and U.S. solid waste management: Effects of the loss of a waste haven. *Journal of the Association of Environmental and Resource Economists 11*(6), 1527–1557.

U.S. Census Bureau (1994). *Geographic Areas Reference Manual*. Washington, DC: U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census.

U.S. Environmental Protection Agency (2020). Advancing sustainable materials management: 2018 tables and figures. Technical report, U.S. Environmental Protection Agency, Office of Resource Conservation and Recovery, Washington, DC.

Viscusi, W. K., J. Huber, and J. Bell (2012). Alternative policies to increase recycling of plastic water bottles in the United States. *Review of Environmental Economics and Policy 6*(2), 190–211.

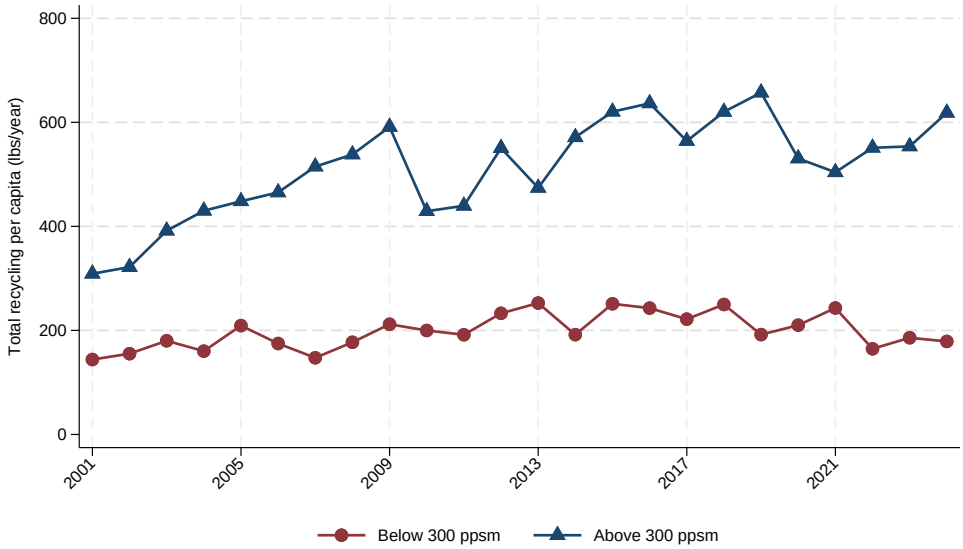
Viscusi, W. K., J. Huber, and J. Bell (2023). Changes in household recycling behavior: Evidence from panel data. *Ecological Economics 208*, 107819.

Viscusi, W. K., J. Huber, J. Bell, and C. Cecot (2013). Discontinuous behavioral responses to recycling laws and plastic water bottle deposits. *American Law and Economics Review 15*(1), 110–155.

Walls, M. (2011). Extended producer responsibility and product design: Economic theory and selected case studies. Discussion Paper RFF DP 11-08, Resources for the Future.

Figures

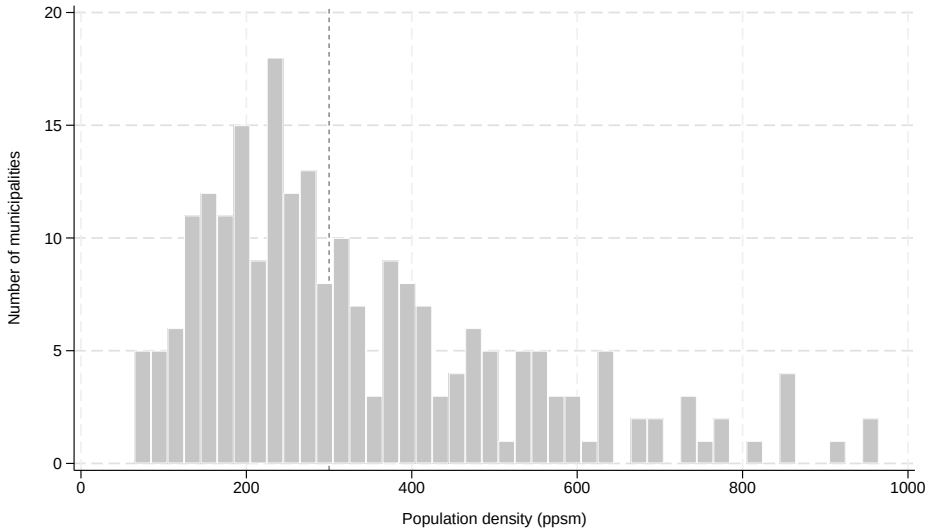
Figure 1: Average total recycling per capita by year for municipalities above and below the 300 ppsm threshold



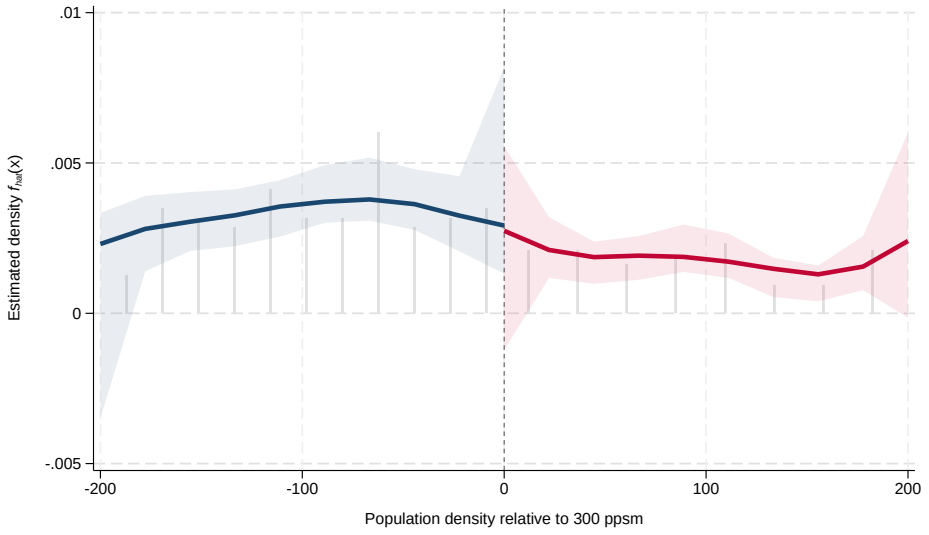
Notes: Average total recycling per capita (lbs/year) by year for municipalities above 300 ppsm (navy triangles) and below 300 ppsm (maroon circles). The figure is restricted to municipalities in the dataset within ±300 of the density threshold. The plot bandwidth $h = 300$ ppsm is chosen for visual legibility; the preferred bandwidth used for estimation is $h = 845$ ppsm (Table 2).

Figure 2: Density of the Running Variable at the 300 ppsm Threshold

(a) Raw density



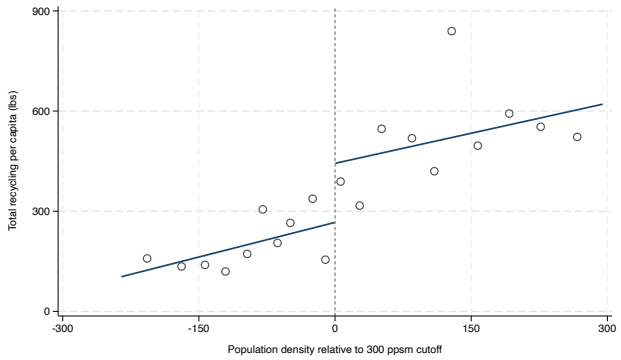
(b) Manipulation test around cutoff



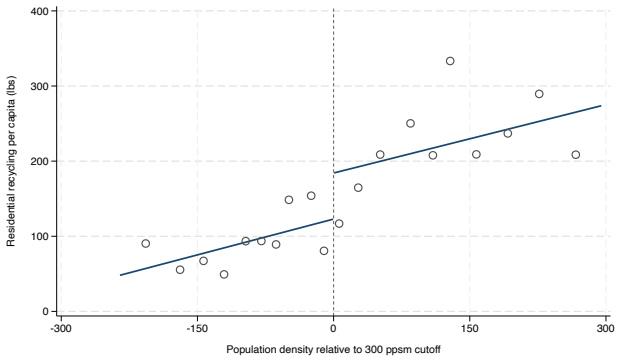
Notes: Cross-section, one observation per municipality at its first Census-era density value ($N = 373$). *Panel (a):* histogram of raw population density (20 ppsm bins, capped at 1,000 ppsm); dashed line marks the 300 ppsm mandate threshold. *Panel (b):* local-polynomial density estimates on the centred running variable $x =$ density $- 300$ over $x \in [-200, 200]$, fitted separately on each side of the cutoff following Cattaneo et al. (2018); shaded bands are 95 percent confidence intervals. Cattaneo–Jansson–Ma manipulation test on the full cross-section: $T = -0.534$, two-sided $p = 0.593$. See Section 4.1 for discussion of the left-tail mass.

Figure 3: Effect of the Recycling Mandate on Reported Recycling at the Density Threshold

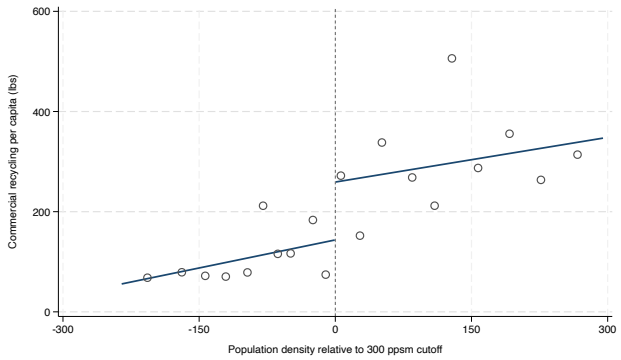
(a) Total recycling



(b) Residential recycling

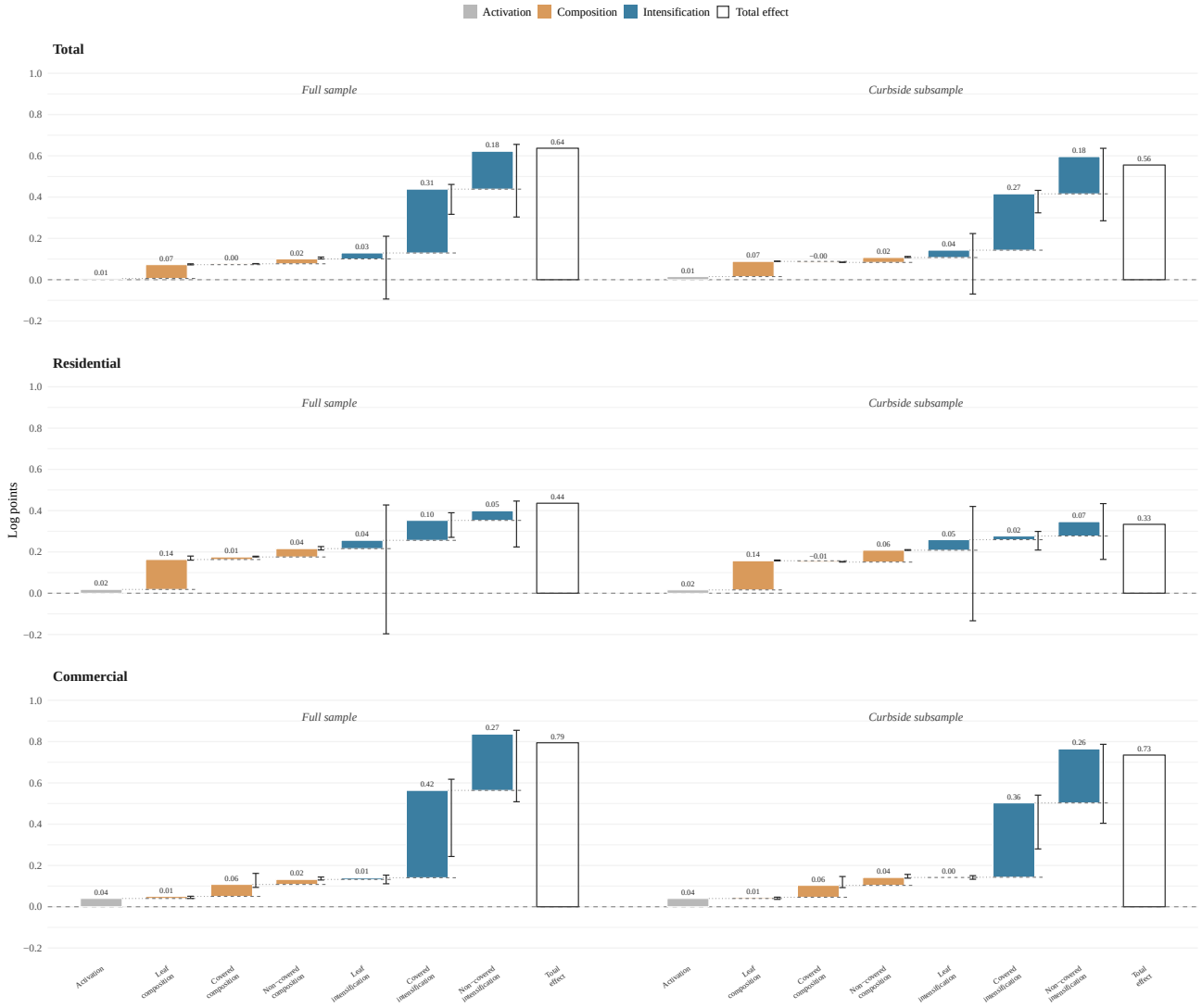


(c) Commercial recycling



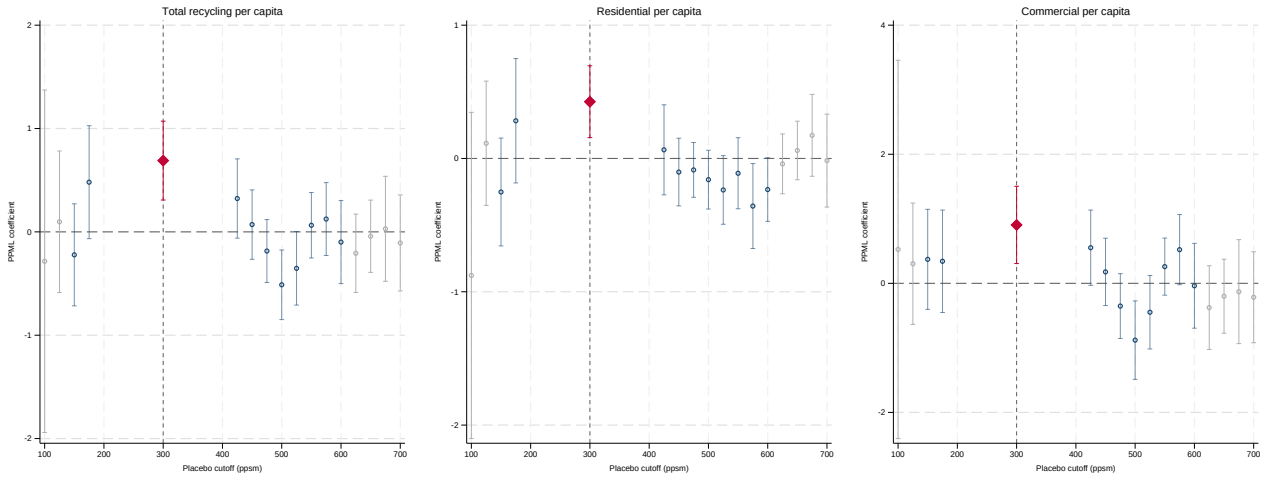
Notes: Each panel plots municipal recycling tonnage per capita (lbs) against population density relative to the statutory threshold. Points represent binned averages within quantile bins of population density relative to the statutory threshold. Lines show separate linear fits on each side of the cutoff. The figure is restricted to municipalities within ± 300 of the density threshold. The plot bandwidth $h = 300$ ppsm is chosen for visual legibility; the preferred bandwidth used for estimation is $h = 845$ ppsm (Table 2).

Figure 4: Decomposing the mandate effect: activation, composition, and intensification by material group



Notes: Decomposition identity of Section 6 on three outcomes (panels) and two samples (Full sample, Curbside subsample). Within each sub-figure, bars sum from left to right to the Total effect (hollow bar, τ): Activation (τ_a); three composition contributions $\omega_k \tau_k^g$ for $k \in \{\text{leaf}, \text{covered}, \text{non-covered}\}$ (ochre); three intensification contributions $\omega_k \tau_k^y$ (blue). Weights ω_k are control-side per-capita-tonnage shares within the RD bandwidth, summing to one across the three groups. Composition and intensification bars carry ω_k-scaled Lee (2009) bounds as dashed whiskers; activation reports a ± 1.96 SE interval. Total-effect bars are hollow to flag that they are the resultant of the contributions, not an additional channel. The covered group for residential and total outcomes contains the eight residential Act 101 materials plus single-stream/commingled aggregates; the covered group for commercial outcomes is Act 101 narrower set (high-grade office paper, corrugated paper, aluminum cans, leaf). Leaf is broken out separately because Act 101 imposes a distinct leaf-collection mandate. Specification: PPML matching Table 2, column (3); curbside subsample restricts to municipalities operating a curbside program in 2001.

Figure 5: Placebo regression discontinuities across the density grid



Notes: Each point is the PPML coefficient on a treatment indicator at an artificial cutoff $c \in [100, 200] \cup [400, 700]$ ppsm in 25 ppsm increments; bars are 95% confidence intervals. The red diamond at $c = 300$ is the true cutoff with bandwidth $h_{300} = 845$ ppsm. Navy points mark the *robust grid* of placebos with at least 30 unique municipalities on each side of c within the local bandwidth; gray points mark sparse placebos that fail this threshold (three on the left, $c \in \{100, 125, 150\}$, and four on the right, $c \in \{625, 650, 675, 700\}$) and are retained in the figure for transparency. Two placebos ($c = 200$, $c = 400$) are infeasible because the one-sided window cannot fit a bandwidth of at least 100 ppsm without reaching the true cutoff, and are omitted. All cells use the benchmark PPML specification in Table 2, column (4). At each placebo the sample is restricted to the side of $c^* = 300$ that contains c (control-only for $c < 300$, treated-only for $c > 300$), and the bandwidth is $h_c = \min\{h_c^{\text{MSE}}, |c - c^*| - 1\}$, floored at 100 ppsm so the window cannot reach the true cutoff (Cattaneo et al., 2020). Across the 48 placebo tests (16 cutoffs $\times$ 3 outcomes), three yield $|t| > 1.96$, versus 2.4 expected under 48 independent 5% tests, and none survives Benjamini–Hochberg adjustment at $q = 0.05$. The true cutoff is the strict maximum of $|t|$ across all 16 placebos for total, residential, and commercial recycling, with randomization-inference $p = 1/17 = 0.059$ on the full grid and $p = 1/11 = 0.091$ on the robust grid. Source: PA DEP administrative records, 2001–2024.

Tables

Table 1: Pre-mandate Covariate Balance at the Density Threshold

	Sample means			Difference		RD estimate	
	Full	Below	Above	Diff.	p	Coeff.	p
White share	0.973	0.985	0.974	-0.011	0.098	0.010	0.688
Prime-age share (18-64)	0.612	0.605	0.619	0.014	0.003	0.009	0.408
College share	0.130	0.103	0.162	0.059	0.000	0.010	0.723
Median income ($000s, 2026)	83.2	81.8	97.1	15.3	0.000	5.1	0.304
Unemployment rate	0.064	0.067	0.055	-0.012	0.001	0.005	0.612
Manufacturing share	0.314	0.313	0.328	0.015	0.259	0.012	0.731
Wholesale + retail trade share	0.190	0.180	0.186	0.006	0.164	-0.007	0.509
FIRE + business services share	0.110	0.095	0.111	0.015	0.001	0.003	0.806
Personal/health/educ services share	0.187	0.169	0.181	0.013	0.141	0.006	0.762
Public administration share	0.038	0.037	0.030	-0.008	0.040	0.008	0.365
Homeownership share	0.754	0.825	0.820	-0.004	0.610	0.013	0.529
Detached-housing share	0.669	0.746	0.776	0.030	0.031	0.023	0.450
Non-control diagnostic:							
First-class township status	0.029	0.000	0.056	0.056	0.013	0.059	0.310

Notes: Columns (1)–(3) report sample means of 1980 (pre-Act 101) Census covariates over the full sample of municipalities with 5,000 to 10,000 inhabitants; the below-threshold subsample, and the above-threshold subsample, all within the baseline bandwidth $h = 845$ ppsm. The table collapses to one observation per municipality (earliest year in the sample); 1980 covariates are time-invariant within municipality. Column (4) is the raw above-minus-below mean difference within the bandwidth; column (5) is its p-value from OLS on a treatment indicator. Columns (6)–(7) are the bias-corrected RD estimate (triangular kernel, $h = 845$) and robust p-value. Standard errors in columns (5) and (7) are clustered at the municipality level. The twelve covariates above the diagnostic row are used in the main regression (column (4) of Table 2); first-class township status is reported as a diagnostic only and is not included in Z_{it}.

Table 2: PPML regression discontinuity estimates at the Act 101 density threshold

	(1)	(2)	(3)	(4)
Panel A: Total recycling per capita				
Above threshold	0.523**	0.513**	0.637***	0.690***
	(0.215)	(0.215)	(0.191)	(0.195)
Mandate effect (%)	68.6	67.0	89.1	99.3
	(36.3)	(35.9)	(36.0)	(38.8)
N (clusters)	230	230	228	228
$N \times t$	3941	3941	3934	3934
Control mean (lbs)	200.0	200.0	200.0	200.0
Panel B: Residential recycling per capita				
Above threshold	0.401**	0.400**	0.436***	0.426***
	(0.197)	(0.198)	(0.144)	(0.137)
Mandate effect (%)	49.3	49.1	54.6	53.1
	(29.5)	(29.5)	(22.2)	(21.0)
N (clusters)	230	230	226	226
$N \times t$	3941	3941	3928	3928
Control mean (lbs)	92.3	92.3	92.3	92.3
Panel C: Commercial recycling per capita				
Above threshold	0.621**	0.604**	0.795***	0.905***
	(0.305)	(0.305)	(0.287)	(0.305)
Mandate effect (%)	86.1	83.0	121.3	147.2
	(56.7)	(55.8)	(63.6)	(75.5)
N (clusters)	230	230	226	226
$N \times t$	3941	3941	3919	3919
Control mean (lbs)	107.7	107.7	107.7	107.7
Year fixed effects	No	Yes	Yes	Yes
County fixed effects	No	No	Yes	Yes
Controls	No	No	No	Yes

Notes: Entries report kernel-weighted PPML regression discontinuity estimates. Column (3) drops municipalities that are singletons in either the county or year fixed-effect dimension; PPML cannot identify these and they are dropped from the likelihood. All columns use a triangular kernel and bandwidth $h = 845$ ppsm, the widest MSE-optimal bandwidth (Calonico et al. 2014) for the three aggregates, selected without covariates to ensure a common estimation sample across streams and specifications. "Above threshold" ($\hat{\tau}$): PPML log-point coefficient on the treatment indicator; municipality-clustered robust standard errors in parentheses; stars on this row only. "Mandate effect (%)": $100 \times (\exp(\hat{\beta}) - 1)$ with delta-method standard error in parentheses. Control mean: raw mean of outcome (lbs per capita) among below-threshold municipalities within the bandwidth. Controls: 1980 pre-Act 101 cross-section — white share, prime-age share, college-educated share, median household income, unemployment rate, owner-occupancy share, single-family detached share, employment shares in manufacturing; wholesale and retail trade; FIRE and business services; personal, health, and educational services; and public administration. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 3: Long-run stability: mandate effects across Census assignment regimes

	(1) 2001–2009	(2) 2010–2019	(3) 2020–2024	(4) p(equal)
Panel A: Total recycling per capita (PPML)				
Above threshold	0.798***	0.845***	0.616**	0.810
	(0.229)	(0.285)	(0.271)	—
Mandate effect (%)	122.1	132.8	85.2	—
	(51.0)	(66.4)	(50.3)	—
N	176	183	170	—
$N \times t$	1390	1740	783	—
Control mean (lbs)	174.3	222.7	197.0	—
Panel B: Residential recycling per capita (PPML)				
Above threshold	0.581***	0.520***	0.363**	0.711
	(0.205)	(0.179)	(0.185)	—
Mandate effect (%)	78.8	68.2	43.7	—
	(36.7)	(30.0)	(26.6)	—
N	176	183	168	—
$N \times t$	1390	1740	781	—
Control mean (lbs)	93.0	88.9	98.9	—
Panel C: Commercial recycling per capita (PPML)				
Above threshold	1.151***	1.116***	0.867**	0.451
	(0.378)	(0.421)	(0.426)	—
Mandate effect (%)	216.2	205.3	138.0	—
	(119.6)	(128.4)	(101.4)	—
N	175	179	168	—
$N \times t$	1385	1718	774	—
Control mean (lbs)	81.3	133.8	98.1	—
Year fixed effects	Yes	Yes	Yes	—
County fixed effects	Yes	Yes	Yes	—
Controls	Yes	Yes	Yes	—

Notes: Panels A–C re-estimate the benchmark full-controls specification of Table 2, column (4), separately by Census assignment regime; entries report the PPML log-point coefficient and $100 \times (\exp(\hat{\beta}) - 1)$ percent effect with delta-method SE. All columns use bandwidth $h = 845$ ppsm (the widest MSE-optimal bandwidth among the three recycling aggregates, estimated on the full panel without covariates). The final column reports the p-value from a Wald test of H_0: $\tau_1 = \tau_2 = \tau_3$, obtained by pooling all three regimes and jointly testing two regime×treatment interactions. The p-value appears on the "Above threshold" row; all other entries in the final column are blank. Standard errors clustered by municipality in parentheses. Stars on the coefficient row only. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 4: Robustness of main PPML estimates

	Total recycling	Residential recycling	Commercial recycling
Panel A: Baseline			
Baseline	0.690***	0.426***	0.905***
	(0.195)	(0.137)	(0.305)
Panel B: Bandwidth variation			
Stream-specific MSE-optimal h_k (768/676/845)	0.658***	0.370***	0.905***
	(0.196)	(0.139)	(0.305)
Bandwidth $h/3$	0.415**	0.147	0.623**
	(0.181)	(0.154)	(0.272)
Bandwidth $h/2$	0.523***	0.272*	0.720**
	(0.197)	(0.145)	(0.308)
Bandwidth $2h$	0.767***	0.495***	0.980***
	(0.197)	(0.135)	(0.311)
CER-optimal bandwidth	0.598***	0.351**	0.799**
	(0.200)	(0.140)	(0.314)
Panel C: Kernel robustness			
Rectangular kernel	0.786***	0.502***	1.011***
	(0.198)	(0.138)	(0.311)
Panel D: Alternative estimator (local-linear CCT, levels)			
Conventional ($\hat{\tau}_{cl}$, % of control mean)	97%***	58%**	131%**
	(34.0%)	(27.1%)	(52.6%)
Bias-corrected ($\hat{\tau}_{bc}$, % of control mean)	79%**	20%	130%**
	(37.7%)	(32.0%)	(58.8%)
Panel E: Donut-hole restrictions			
Donut hole ±10	0.637***	0.433***	0.807**
	(0.221)	(0.150)	(0.349)
Donut hole ±20	0.649**	0.458***	0.818**
	(0.254)	(0.162)	(0.406)
Donut hole ±30	0.802**	0.623***	0.938*
	(0.344)	(0.231)	(0.511)

(continued on next page)

Table 4: Robustness of main PPML estimates *(continued from previous page)*

	Total recycling	Residential recycling	Commercial recycling
Panel F: Sample restriction			
Excl. first-class townships	0.716***	0.439***	0.956***
	(0.198)	(0.138)	(0.312)
Panel G: County-clustered standard errors			
County-clustered SE	0.690***	0.426***	0.905**
	(0.237)	(0.111)	(0.363)
Panel H: Conley (1999) spatial SEs			
Conley SE (25 km cutoff)	0.690***	0.426***	0.905***
	(0.166)	(0.091)	(0.267)
Conley SE (50 km cutoff)	0.690***	0.426***	0.905***
	(0.184)	(0.073)	(0.292)
Year fixed effects	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Baseline bandwidth h	845	845	845

Notes: Each cell reports the PPML coefficient $\hat{\beta}$ on the treatment indicator from the indicated specification; entries in parentheses are standard errors clustered at the municipality unless otherwise noted. The baseline (Panel A) replicates the benchmark full-controls specification in Table 2, column (4). Panel B varies the bandwidth: outcome-specific MSE-optimal ($h_{\text{tot}} = 768$, $h_{\text{res}} = 676$, $h_{\text{com}} = 845$ ppsm) and common $h/3$, $h/2$, $2h$, and CER-optimal. Panel C replaces the triangular kernel with a rectangular (uniform) kernel. Panel D replaces PPML with the local-linear estimator of Calonico et al. (2014), with year and county dummies entered as additive covariates; because the local-linear coefficient is in levels, entries are rescaled as a percent of the control-side mean within the bandwidth for direct comparability with the implied PPML percent effects $(\exp(\hat{\beta}) - 1) \cdot 100$. Panel E imposes donut-hole exclusions of ± 10, ± 20, and ± 30 ppsm around the cutoff. Panel F excludes the 91 municipalities that held first-class township status when Act 101 was enacted in 1988 (the indicator used is `is_first_class_twp_1988`); the baseline bandwidth is retained. Panels G–H address inference robustness with point estimates identical to Panel A by construction: Panel G re-clusters standard errors at the county level; Panel H reports Conley (1999) spatial HAC standard errors with Bartlett kernel at 25 km and 50 km cutoffs. All PPML specifications use the preferred twelve 1980 controls and $h = 845$ ppsm as the base unless otherwise noted. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Contents

A Data and Measurement **39**
A.1 Geographic Distribution and Sample-Construction Funnel 39
A.2 Covariate Balance for Excluded Non-Reporters 39

B RD Design and Robustness **39**
B.1 Controls Robustness 40
B.2 Kappa-weighted characterization of the local RD population 40
B.3 Placebo cutoffs 40

C Decomposition: Algebra, Bounds, and Robustness **40**
C.1 Decomposition algebra 41
C.2 Residual term 43
C.3 Lee bounds for selected margins 43
C.4 Why the monotonicity assumption is plausible here 44
C.5 Specification used for the bounds and covariate-adjusted sensitivity 46
C.6 Main decomposition tables 46

A Data and Measurement

This appendix documents the construction of the analysis sample used throughout Sections 3 to 5 and gathers a robustness check on reporting-related sample selection. Appendix A.1 illustrates the geographic distribution of the design sample and reports the sample-construction funnel from the Census universe to the preferred RD bandwidth. Appendix A.2 re-runs the covariate-balance test on the 527 municipality-years excluded from the outcome analysis because they lack individual tonnage, showing that the reporting-induced sample restriction is balanced across the 300 ppsm cutoff.

A.1 Geographic Distribution and Sample-Construction Funnel

Figure A1 maps the 2020 cross-sectional classification of Pennsylvania municipalities into in-sample treated, in-sample control, and out-of-sample cells. Table A1 reports the cumulative restrictions that take the panel from the full Census universe of Pennsylvania municipality-years to the preferred RD bandwidth.

Each row of Table A1 applies its restriction cumulatively to the prior row. Row 1 is the full U.S. Census Gazetteer universe of Pennsylvania municipality-years, 2001–2024. Row 2 restricts to municipalities with population 5,000–10,000 under the prevailing decennial benchmark, the range over which the 300 ppsm density rule is the binding Act 101 assignment margin. Row 3 retains municipality-years with an individual tonnage report in the DEP system (excluding county-aggregate-only and missing records). Row 4 drops municipality-years flagged as data-quality outliers (per capita tonnage thresholds defined in Section 3); the filter operates within municipality, so a row can lose municipality-years without losing municipalities (e.g., the outlier filter drops 117 flagged municipality-years from 44 municipalities, but those 44 also have non-flagged years and so remain in the sample). Row 5 restricts to the common preferred RD bandwidth $h = 845$ ppsm — the largest of the three stream-specific MSE-optimal bandwidths (commercial), pooled across outcomes to keep cross-stream comparisons on a common identifying sample — and defines the main estimation sample. The "Municipality-years" column counts observations; the "Municipalities" column counts unique municipalities ever appearing in the row.

A.2 Covariate Balance for Excluded Non-Reporters

Section 3 excludes 527 municipality-years that lack individual recycling tonnage from the outcome analysis. If this exclusion were differential across the 300 ppsm threshold, the recycling discontinuity could reflect selection into reporting rather than program response. Table A2 addresses this concern by re-running the covariate balance check on the excluded subsample. Across all twelve demographic and economic covariates, the RD coefficient is statistically indistinguishable from zero ($p > 0.10$ in every row), indicating that municipalities dropping out of the reporting sample do not differ systematically across the cutoff. Combined with the formal RD test on the probability of individual reporting (Section 3, footnote), this rules out differential selection as a confound for the main RD estimates.

B RD Design and Robustness

This appendix collects diagnostics for the regression-discontinuity design of Section 4 and robustness checks on the main RD estimates in Section 5. It is organized around three checks.

Appendix B.1 tests whether the 1980 pre-Act controls move the preferred coefficient, via a Wald equality test across nested control sets. Appendix B.2 characterizes the local RD population using kernel (κ) weights, to clarify which municipalities receive the greatest weight in the preferred estimate and how they compare to the equal-weighted Pennsylvania 5,000–10,000 population sample and the national portability target. Appendix B.3 examines whether the main RD jump is specific to the 300-ppsm statutory threshold by re-estimating the design at a grid of placebo cutoffs and ranking the true cutoff against the placebo distribution.

B.1 Controls Robustness

This appendix tests whether the 1980 pre-Act controls move the benchmark RD point estimate. Table B1 reports a Wald test of equality of the treatment coefficient across five nested control sets — no controls, Demographics, Human capital & income, Industry composition, Housing, added cumulatively — estimated as a stacked PPML in which observations are replicated five times and the treatment, running-variable slope, controls, and absorbed fixed effects (year, county) are each interacted with the control-set indicator. Standard errors are clustered at the municipality level. The FE structure is held fixed at the benchmark column (4) of Table 2; only the control set varies.

B.2 Kappa-weighted characterization of the local RD population

Because the design is sharp, there is no separate compliance margin; the kernel weights instead identify which municipalities the RD estimate most heavily reflects. Table B2 reports baseline characteristics under three concentric scopes: the kernel-weighted local sample inside the preferred bandwidth, the equal-weighted Pennsylvania 5,000–10,000 population sample, and the corresponding U.S. 5,000–10,000 population sample excluding Pennsylvania. The comparison documents how representative the local RD population is of the broader Pennsylvania population-window sample that enters the regression before kernel weighting, and of the national portability target.

B.3 Placebo cutoffs

Table B3 reports per-cutoff diagnostics (sample sizes, point estimates, t-statistics) that complement the visual evidence in Figure 5. The rank-based p-values obtained by ranking the true cutoff's $|t|$ against the placebo distribution are $1/17 = 0.059$ on the full grid of 16 feasible placebos and $1/11 = 0.091$ on the sample-size-restricted "robust" grid (already reported in the notes to Figure 5). These are diagnostic rather than formal randomization-inference p-values, since placebo windows are side-restricted and use smaller realized bandwidths than the true-cutoff window.

C Decomposition: Algebra, Bounds, and Robustness

This appendix supports the mechanism analysis of Section 6. Appendices C.1 to C.5 derive the activation–composition–intensification decomposition on the material partition $\mathcal{K}_3 = \{\ell, m, n\}$ used in the main text, state the Lee-bounds construction used for the selected margins, defend the monotonicity assumption, and document the bounds specification together with the covariate-adjusted sensitivity. Appendix C.6 collects the headline decomposition tables, including the baseline-curbside subsample re-estimate.

Notation follows the main text. The decomposition conditions on the sample of municipality-years with observed individual municipal tonnage for the outcome; non-individual cases (county-aggregate and no-record municipality-years) carry missing tonnage and are excluded, with Section 3 separately documenting that their share is small and smooth at the cutoff. Within this sample, a_{it} is an indicator for positive reported tonnage on the outcome (activation); for each $k \in \mathcal{K}_3 = \{\ell, m, n\}$, g_{kit} indicates whether material group k appears with positive tonnage among activated municipality-years, and y_{kit} is per capita tonnage within group k when $g_{kit} = 1$. The activation margin defined here is distinct from Section 3's individual-reporting coverage diagnostic: that diagnostic compares individual filings to county-aggregate and no-record cases on the full 5,000–10,000 population sample, while activation is a within-sample zero-to-positive jump. The material groups are defined by Act 101: ℓ is leaf waste, which the statute covers through a separate yard-waste mandate; m collects the remaining covered materials—paper, glass, metal cans, and eligible plastics in the residential outcome, and the narrower §1501(c)(iii) list (high-grade office paper, corrugated paper, aluminum cans) in the commercial outcome; and n collects all other reported categories. Single-stream and commingled aggregates are assigned to m because the DEP aggregate categories primarily bundle these curbside recyclable materials rather than separately identifying their statutory components. For the total outcome, the covered set $\ell \cup m$ is the per-outcome sum.

C.1 Decomposition algebra

Throughout this subsection, all expectations, limits, and identities condition on the Section 6 estimation sample: municipality-years with observed individual municipal tonnage for the outcome. On municipality-years with $a_{it} = 0$, reported tonnage is zero by construction; on municipality-years with $a_{it} = 1$, reported tonnage is the sum of leaf, non-leaf covered, and non-covered tonnage. Combining these two cases,

$$y_{it} = a_{it} \sum_{k \in \mathcal{K}_3} g_{kit}\, y_{kit} \quad \text{a.s. (conditional on the sample).} \tag{4}$$

Taking conditional expectations given the running variable x and mandate status T, and applying the law of iterated expectations,

$$E[y_{it} \mid x, T] = \Pr(a_{it} = 1 \mid x, T) \cdot E\left[\sum_{k \in \mathcal{K}_3} g_{kit} y_{kit} \,\middle|\, a_{it} = 1, x, T \right]. \tag{5}$$

Define each material group's contribution among activated municipality-years as

$$\mu_k(x, T) \equiv E[g_{kit} y_{kit} \mid a_{it} = 1, x, T], \qquad k \in \mathcal{K}_3.$$

Then

$$\log E[y_{it} \mid x, T] = \log \Pr(a_{it} = 1 \mid x, T) + \log\left(\sum_{k \in \mathcal{K}_3} \mu_k(x, T) \right). \tag{6}$$

Each material-group contribution $\mu_k(x, T)$ in Equation (6) can itself be factored into a composition margin and an intensification margin among activated municipality-years. For $k \in \mathcal{K}_3$,

$$\mu_k(x, T) = E[g_{kit} y_{kit} \mid a_{it} = 1, x, T]$$

$$= \Pr(g_{kit} = 1 \mid a_{it} = 1, x, T) \cdot E[y_{kit} \mid g_{kit} = 1, a_{it} = 1, x, T]. \tag{7}$$

The first factor is the probability that material group k appears with positive tonnage among activated municipality-years; the second is per capita tonnage conditional on that group appearing. Taking logs of Equation (7) gives

$$\log \mu_k(x, T) = \log \Pr(g_{kit} = 1 \mid a_{it} = 1, x, T) + \log E[y_{kit} \mid g_{kit} = 1, a_{it} = 1, x, T]. \tag{8}$$

Equation (6) leaves one aggregation step: expressing the log-sum $\log \sum_k \mu_k$ in terms of the separately decomposed material-group contributions $\log \mu_k$ from (8). To do so, expand the log-sum around the corresponding control-side activated means

$$\mu_k^0 \equiv E_h[g_{kit} y_{kit} \mid T = 0, a_{it} = 1, |x| \le h], \qquad k \in \mathcal{K}_3.$$

Here E_h denotes the mean over municipality-years inside the RD bandwidth, pooled across sample years. Let

$$\omega_k \equiv \frac{\mu_k^0}{\sum_{j \in \mathcal{K}_3} \mu_j^0}, \qquad k \in \mathcal{K}_3, \qquad \omega_\ell + \omega_m + \omega_n = 1. \tag{9}$$

Thus ω_k is the share of total per capita recyclables accounted for by material group k among control-side activated municipality-years inside the bandwidth, pooling all sample years. These are the baseline material shares used in Section 6. The combined covered share is $\omega_c \equiv \omega_\ell + \omega_m$.

A first-order expansion of $\log \sum_k \mu_k$ in log deviations gives

$$\log\left(\sum_{k \in \mathcal{K}_3} \mu_k(x, T)\right) = \text{constant} + \sum_{k \in \mathcal{K}_3} \omega_k \log \mu_k(x, T) + \varepsilon_{\text{Jensen}}(x, T), \tag{10}$$

where the residual is zero when all three material groups have the same log change and is second order in the dispersion of their log changes.

Combining (6), (8), and (10), and applying the RD jump operator

$$\Delta f \equiv \lim_{x \downarrow 0} f(x) - \lim_{x \uparrow 0} f(x),$$

yields the decomposition reported as Equation (3) in the main text:

$$\tau = \underbrace{\tau_a}_{\text{activation}} + \underbrace{\sum_{k \in \mathcal{K}_3} \omega_k \tau_k^g}_{\text{composition}} + \underbrace{\sum_{k \in \mathcal{K}_3} \omega_k \tau_k^y}_{\text{intensification}} + \varepsilon \tag{11}$$

The components are

$$\begin{aligned}
&\tau_a \equiv \Delta \log \Pr(a_{it} = 1 \mid x), \\
&\tau_k^g \equiv \Delta \log \Pr(g_{kit} = 1 \mid a_{it} = 1, x), && k \in \{\ell, m, n\}, \\
&\tau_k^y \equiv \Delta \log E[y_{kit} \mid g_{kit} = 1, a_{it} = 1, x], && k \in \{\ell, m, n\}.
\end{aligned}$$

In these component definitions, mandate status is implicit in the one-sided limits defining Δ. The residual ε contains the RD jump in $\varepsilon_{\text{Jensen}}(x, T)$ from the first-order log-sum expansion; the next subsection decomposes the empirical reconstruction residual into this curvature term and the remaining implementation wedges. Define the total per-group log-jump $\Delta_k \equiv \tau_k^g + \tau_k^y$ for use in the curvature analysis below.

C.2 Residual term

The reconstruction residual ε in (11) is decomposed into three pieces. First, the log-sum expansion is not exact unless all three material groups have the same log change. With $\Delta_k \equiv \tau_k^g + \tau_k^y$ for $k \in \mathcal{K}_3$, the exact between-group curvature term is

$$\varepsilon_{\text{Jensen}} = \log\left(\sum_{k\in\mathcal{K}_3} \omega_k e^{\Delta_k}\right) - \sum_{k\in\mathcal{K}_3} \omega_k \Delta_k. \tag{12}$$

By Jensen's inequality, $\varepsilon_{\text{Jensen}} \geq 0$, with equality if and only if $\Delta_\ell = \Delta_m = \Delta_n$. A second-order expansion gives the standard pairwise-variance approximation

$$\varepsilon_{\text{Jensen}} \approx \frac{1}{2}\sum_{j<k} \omega_j \omega_k \left(\Delta_j - \Delta_k\right)^2, \tag{13}$$

which is exact through $O(\|\Delta\|^2)$ around a common mean. Equation (13) makes explicit that the curvature term loads on the three pairwise gaps $\Delta_\ell - \Delta_m$, $\Delta_\ell - \Delta_n$, and $\Delta_m - \Delta_n$, weighted by the corresponding pairs of baseline shares.

Second, the empirical implementation estimates the baseline-share weights from the control-side filing cell inside the bandwidth. The numerical diagnostic reports the small wedge from replacing the table weights ω_k with their bootstrap means $\bar{\omega}_k$:

$$\varepsilon_\omega = \sum_{k\in\mathcal{K}_3} \left(\bar{\omega}_k - \omega_k\right) \Delta_k. \tag{14}$$

Because both ω_k and $\bar{\omega}_k$ sum to one, ε_ω depends only on dispersion in the Δ_k across groups; if all Δ_k are equal, $\varepsilon_\omega = 0$.

Third, after accounting for log-sum curvature and the weight-estimation wedge, any remaining discrepancy reflects the fact that the empirical components are estimated separately and on nested conditional samples. This remaining reconstruction term is collected in ρ:

$$\rho \equiv \tau - \tau_a - \sum_{k\in\mathcal{K}_3} \omega_k(\tau_k^g + \tau_k^y) - \varepsilon_{\text{Jensen}} - \varepsilon_\omega. \tag{15}$$

Thus

$$\varepsilon = \varepsilon_{\text{Jensen}} + \varepsilon_\omega + \rho.$$

Table C1 reports this decomposition numerically. The exact Jensen residual $\varepsilon_{\text{Jensen}}$, the weight-estimation wedge ε_ω, and the remaining reconstruction term ρ are each small relative to the reduced-form log-jump on every outcome, and the second-order Jensen approximation in (13) tracks the exact curvature term closely.

C.3 Lee bounds for selected margins

The activation margin τ_a is point-identified as a cutoff jump in the probability of positive tonnage on the Section 6 sample. The remaining six margins—three composition margins τ_k^g and three intensification margins τ_k^y, $k \in \mathcal{K}_3$—are conditional on treatment-affected selected samples. I bound them using the one-sided selection logic of Lee (2009).

Assumption 1 (One-sided selection)**.** *At the cutoff,*

$$a_{it}(1) \geq a_{it}(0) \quad a.s., \tag{16}$$
$$g_{\ell it}(1) \geq g_{\ell it}(0) \quad a.s., \tag{17}$$
$$g_{mit}(1) \geq g_{mit}(0) \quad a.s., \tag{18}$$
$$g_{nit}(1) \geq g_{nit}(0) \quad a.s.. \tag{19}$$

Assumption 1 rules out mandate-induced exit from the relevant selected sample. Equation (16) rules out the pattern in which a municipality reports positive tonnage when unmandated yet none when mandated. Equations (17) to (19) rule out the analogous pattern for the leaf, non-leaf covered, and non-covered material groups, respectively: positive group-k tonnage when unmandated yet none when mandated.

Composition margins. For τ_k^g, $k \in \mathcal{K}_3$, the outcome is g_{kit} and the selection event is activation, $a_{it} = 1$. The trim share is the level selection differential

$$q_a \equiv \frac{\Pr(a_{it} = 1 \mid x = 0^+, T = 1) - \Pr(a_{it} = 1 \mid x = 0^-, T = 0)}{\Pr(a_{it} = 1 \mid x = 0^+, T = 1)}. \tag{20}$$

The same q_a applies to all three composition margins because each is selected through the same activation event. Trimming the appropriate tail of the treated-side distribution of g_{kit} yields Lee lower and upper endpoints for τ_k^g, $k \in \mathcal{K}_3$.

Intensification margins. For τ_k^y, $k \in \mathcal{K}_3$, the outcome is y_{kit} and the selection event is material-group appearance among activated municipality-years. The group-specific trim share is

$$q_C^k \equiv \frac{\Pr(g_{kit} = 1 \mid x = 0^+, T = 1, a_{it} = 1) - \Pr(g_{kit} = 1 \mid x = 0^-, T = 0, a_{it} = 1)}{\Pr(g_{kit} = 1 \mid x = 0^+, T = 1, a_{it} = 1)}, \qquad k \in \mathcal{K}_3. \tag{21}$$

Trimming the treated-side conditional distribution of y_{kit} gives Lee lower and upper endpoints for τ_k^y.

The bounds are sharp at each selection stage under Assumption 1 and the local independence condition underlying the RD design. I do not impose joint sharpness across all six selected margins. When aggregating bounded components into the partition-aggregated composition rectangle $\sum_k \omega_k \tau_k^g$ or the partition-aggregated intensification rectangle $\sum_k \omega_k \tau_k^y$, I therefore sum lower endpoints and upper endpoints separately across the three groups. This is conservative: it allows the worst case in each selected margin to occur simultaneously across ℓ, m, and n.

C.4 Why the monotonicity assumption is plausible here

The four inequalities in Assumption 1 have different institutional status.

For activation and for the two covered-material composition margins (g_ℓ and g_m), the restriction is grounded in the statute. Act 101 imposes a one-sided legal regime on municipalities above the cutoff. §1501(a) and §1501(b) impose the municipal recycling-program duty and enumerate the residential non-leaf covered materials; §1501(d) separately imposes the leaf-waste collection obligation; §1501(c)(iii) imposes a narrower covered-material obligation on commercial, municipal,

and institutional generators within mandated municipalities; §1707 gives DEP authority to require records and reports; and §904(b) ties annual grant disbursement to source-separated tonnage submissions. These provisions add affirmative collection and reporting obligations on the mandated side; they do not remove them. There is no statutory mechanism by which mandate exposure would make a municipality, within the §6 sample of observed individual municipal tonnage, less likely to report positive tonnage on the outcome, in the leaf group, or in the non-leaf covered group than its unmandated counterfactual. The leaf inequality (17) is reinforced by the institutional fact that leaf-waste collection programs predated Act 101 in many municipalities and operate under a distinct yard-waste mandate after enactment; there is no plausible behavioral channel by which mandate exposure would induce a municipality with a pre-existing leaf program to stop reporting leaf tonnage.

The non-covered restriction (19) is weaker. Act 101 neither compels nor prohibits reporting of non-covered materials. A municipality could, in principle, reallocate reporting effort toward statutorily named materials and away from discretionary categories. Such behavior would violate (19). For this reason, the non-covered Lee bounds should be read as relying on an additional behavioral restriction. In the estimates, the non-covered bounds are wide and often straddle zero, so relaxing (19) would weaken claims about the non-covered response but would not overturn the non-leaf covered intensification result.

Assumption 1 is cross-sectional. It compares potential outcomes for the same municipality-year at the cutoff under mandate and no mandate. It does not require activation or material-group appearance to be monotone over time within a municipality. Year-over-year gaps in the panel therefore do not directly test the assumption. For transparency, the δ-sensitivity exercise below perturbs the Lee trim shares.

δ-sensitivity. The sensitivity exercise perturbs the trim share separately at each Lee step. For a composition margin τ_k^g the relevant trim share is the activation-selection share q_a from (20); for an intensification margin τ_k^y the relevant trim share is the group-k selection share q_C^k from (21). Let

$$\tau_k^{g,L}(q),\ \tau_k^{g,U}(q)$$

denote the Lee lower and upper endpoints for τ_k^g when the treated-side activated sample is trimmed by share q, and let

$$\tau_k^{y,L}(q),\ \tau_k^{y,U}(q)$$

denote the corresponding endpoints for τ_k^y when the treated-side group-k sample among activated municipality-years is trimmed by share q.

To allow a fraction δ of treated observations to violate monotonicity, the lower endpoint uses a larger trim share and the upper endpoint uses a smaller trim share. Thus the sensitivity bounds for group-k intensification are

$$\left[\tau_k^{y,L}(q_C^k+\delta),\quad \tau_k^{y,U}\big(\max\{q_C^k-\delta,0\}\big)\right],\qquad k\in\mathcal{K}_3.$$

I evaluate these bounds for $\delta \in \{0, 0.01, 0.025, 0.05\}$ in Table C2. The load-bearing object is τ_m^y, the non-leaf covered intensification jump, since the section's main quantitative claims concern that margin in the total and commercial streams. Its lower endpoint $\tau_m^{y,L}$ remains strictly positive at

$\delta = 0.05$ for total ($\tau_m^{y,L} = 0.229$) and commercial ($\tau_m^{y,L} = 0.062$) recycling, so the headline covered-material intensification result is not overturned by allowing five percent of treated observations in the relevant cell to violate monotonicity. For residential recycling the point-identified bound at $\delta = 0$ is already narrow, and $\tau_m^{y,L}$ crosses zero between $\delta = 0.025$ and $\delta = 0.05$; this is consistent with the main-text reading that the residential covered-intensification response is more fragile and weakens further in the curbside subsample. The leaf and non-covered intensification bounds are wide at $\delta = 0$ and remain so under perturbation, reflecting the larger selection differentials in those groups; the paper does not draw causal claims from those margins.

C.5 Specification used for the bounds and covariate-adjusted sensitivity

The decomposition bounds are estimated using the column-(3) specification of Table 2: separate slopes, a triangular kernel, year and county fixed effects, and no continuous covariates. This specification preserves the Lee trim shares as level selection-rate gaps. Year and county fixed effects difference out group means without changing the level selection-rate gap used to define the trim, whereas continuous covariates in the PPML index would change the object being trimmed after partialling out. A Lee-compatible adjustment for continuous covariates would require stratifying on covariates and aggregating cell-level bounds (Lee, 2009, §3), which is infeasible with twelve continuous covariates and the sample size available.

As a sensitivity check, Table C3 reports the corresponding point estimates after adding the twelve 1980 pre-Act covariates to the PPML index. The treatment coefficients are very close to the main estimates in Table C4, reported also in column (1) of that table. Because the point-identified margins are essentially unchanged, there is little empirical reason to expect the Lee bounds to change substantively under a feasible covariate-adjusted implementation.

C.6 Main decomposition tables

This subsection collects the decomposition tables that underlie the narrative in Section 6. The partition $\mathcal{K}_3 = \{\ell, m, n\}$ separates leaf waste ℓ from the remaining non-leaf covered materials m, leaving non-covered tonnage n as the third group; the combined covered share is $\omega_c = \omega_\ell + \omega_m$. The identity is (11), restated for reference:

$$\tau = \tau_a + \sum_{k \in \mathcal{K}_3} \omega_k \tau_k^g + \sum_{k \in \mathcal{K}_3} \omega_k \tau_k^y + \varepsilon$$

with the same control-side kernel-weighted weights as in the main text.

Table C4 reports the full-sample decomposition for total, residential, and commercial recycling. This is the source for the headline activation, composition, intensification, and per-channel contribution numbers used in Section 6.

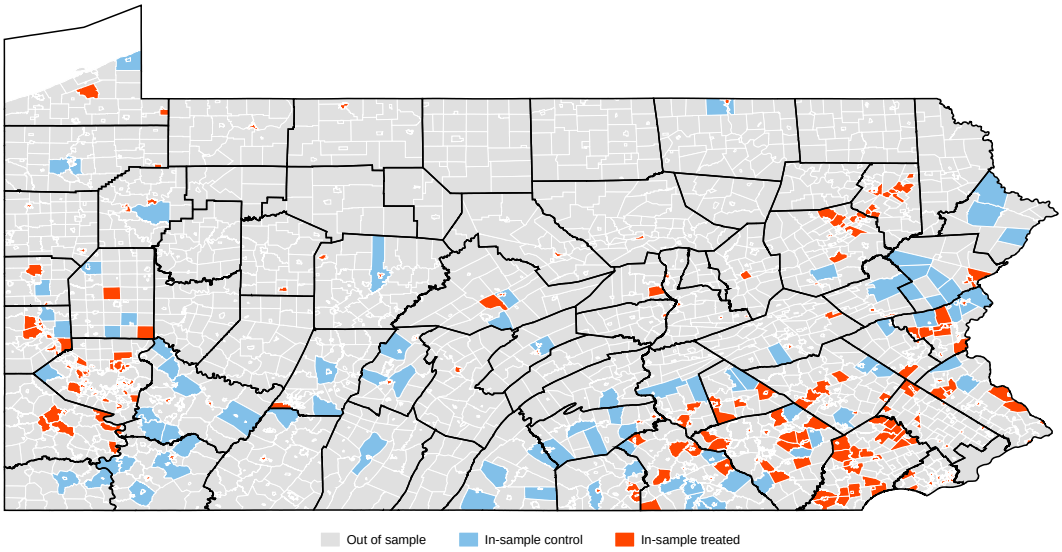
The curbside-subsample check uses two linked pieces of evidence. Because pre-Act 101 curbside status is unobserved, the subsample is defined by curbside operation in 2001, the first year of the outcome panel. This is a post-enactment restriction, not a pre-treatment "already curbside" sample. DEP reports show that curbside programs were limited before Act 101 and expanded substantially during the post-enactment period: the number of municipalities with curbside programs rose from 68 in April 1986 to 146 in December 1988, 562 by September 1991, 697 curbside-only municipalities in 1997 (879 including municipalities with both curbside and drop-off service), and

945 annual-report curbside programs in 2001 (Pennsylvania Department of Environmental Resources, 1991b, Table 3, p. 5; Pennsylvania Department of Environmental Resources, 1991a, p. 2; Pennsylvania Department of Environmental Protection, 1998, Table 2, p. 1; Pennsylvania Department of Environmental Protection, 2001, Table 1, p. 1). DEP annual-report totals and the project registry do not use identical counting rules, but both show that the 2001 curbside measure is far downstream of Act 101 and the post-enactment diffusion of curbside service. The project panel's 2001 curbside subsample contains 1,012 municipalities with curbside service recorded at the start of the 2001–2024 outcome panel.

For the restriction to be informative as an interpretive check, the conditioning variable should itself be smooth at the cutoff. Table C5 reports a sharp RD on the 2001 cross-section: the discontinuity in baseline curbside operation is not statistically distinguishable from zero (robust $p = 0.30$ without controls; $p = 0.48$ with the 1980 controls), so the subsample restriction is not selecting on a variable that jumps at the threshold. Table C6 then re-runs the decomposition on the same 2001-curbside subsample. The exercise asks whether the full-sample pattern is mechanically driven by municipalities with any recorded curbside program at the start of the panel. It does not hold fixed the full Act 101 compliance bundle: the registry records the existence of a curbside program, not whether the program satisfies the statute's collection, education, enforcement, or reporting requirements, or how intensively covered materials are collected and recorded.

Appendix Figures

Figure A1: Geographic Distribution of the 2020 Design Sample.



Notes: The map classifies Pennsylvania municipalities using 2020 Census population and density values. In-sample treated municipalities have population between 5,000 and 10,000 and density of at least 300 persons per square mile. In-sample control municipalities have population between 5,000 and 10,000 and density below 300. All remaining municipalities are out of sample for the main regression discontinuity design. The empirical panel updates this assignment at each decennial Census; the map is a 2020 cross-sectional illustration.

Appendix Tables

Table A1: Sample Construction Funnel

Sample	Municipality-years	Municipalities	Share of prior row
Full Gazetteer universe (2001–2024)	61,811	2,587	—
In-band (5,000–10,000 population)	7,521	373	12.2%
With individual tonnage report	6,994	371	93.0%
After outlier exclusion	6,877	371	98.3%
Within RD bandwidth ($h = 845$ ppsm)	3,941	230	57.3%

Notes: Each row applies the indicated restriction cumulatively to the prior row. "Municipality-years" counts observations; "Municipalities" counts unique municipalities ever in the row's sample. See Appendix A.1 for row-by-row definitions, the outlier filter, and the bandwidth choice $h = 845$ ppsm.

Table A2: Covariate balance at the 300 ppsm threshold: excluded non-reporters

	Sample means			Difference		RD estimate	
	Full	Below	Above	Diff.	p	Coeff.	p
White share, 1980	0.982	0.988	0.989	0.001	0.870	0.005	0.518
Prime-age share (18-64), 1980	0.603	0.601	0.617	0.016	0.058	0.018	0.207
College share, 1980	0.104	0.092	0.124	0.032	0.118	-0.006	0.922
Median income, 1980 ($000s, 2026)	79.1	80.1	83.7	3.6	0.275	1.2	0.871
Unemployment rate, 1980	0.073	0.070	0.073	0.003	0.656	0.024	0.174
Manufacturing share, 1980	0.319	0.317	0.347	0.030	0.276	-0.031	0.695
Wholesale + retail trade share, 1980	0.190	0.182	0.192	0.010	0.281	-0.028	0.290
FIRE + business services share, 1980	0.101	0.095	0.096	0.001	0.873	0.011	0.683
Personal/health/educ services share, 1980	0.173	0.161	0.182	0.022	0.237	0.005	0.887
Public administration share, 1980	0.040	0.037	0.035	-0.002	0.768	0.010	0.544
Homeownership share, 1980	0.781	0.821	0.806	-0.015	0.410	0.033	0.279
Detached-housing share, 1980	0.717	0.750	0.739	-0.011	0.664	0.052	0.266

Notes: Sample restricted to the 527 municipality-years excluded from the outcome analysis because they lack individual recycling tonnage in the DEP database (county-level aggregate entries or no DEP record). "Below" and "Above" columns report unweighted means below and above the 300 ppsm cutoff. "Diff." is the raw mean difference with the corresponding two-sided p-value. "RD estimate" is the local-linear coefficient at the cutoff with a triangular kernel and the same MSE-optimal bandwidth used in the main analysis; standard errors clustered by municipality. U.S. Census/ACS demographic covariates merged to PA DEP municipality panel, 2001–2024.

Table B1: Equality of the preferred coefficient across nested control sets

Outcome	Joint equality p (cols 0–4)	Col (0) vs. Col (4) p
Total recycling	0.674	0.330
Residential	0.642	0.846
Commercial	0.556	0.229

Notes: Wald test of equality of the log-scale treatment coefficient across five nested 1980-control sets (no controls; +Demographics; +Human capital & income; +Industry composition; +Housing), estimated as a stacked PPML in which observations are replicated five times (one per control set) and the treatment, running-variable slope, controls, and absorbed fixed effects (year, county) are each interacted with the control-set indicator. Standard errors clustered at municipality level. The joint column reports the four-degree-of-freedom test that the treatment coefficients in the four control-augmented specifications all equal the no-controls baseline; the pairwise column reports the one-degree-of-freedom test of the no-controls baseline vs. the full twelve-control specification.

Table B2: Kappa-weighted characterization of the local RD population

	Local RD		PA 5–10k		US 5–10k excl. PA	
	Mean	SD	Mean	SD	Mean	SD
Median household income (USD)	101,052	31,995	90,471	34,246	80,923	38,166
Bachelor's or higher (%)	30.6	13.8	30.4	14.2	30.6	15.9
White-alone (%)	89.4	6.6	84.9	12.0	81.3	17.5
Unemployment rate (%)	3.9	2.4	4.7	2.7	4.8	3.0
Median age (years)	—	—	—	—	41.8	6.5
Municipalities	180		306		1128	

Notes: The three column groups are concentric scopes around the 300 ppsm Act 101 density cutoff. *Local RD* is the Pennsylvania 2020 cross-section within the preferred bandwidth $h = 845$ ppsm, weighted by the triangular kernel used in the RD estimation. Because the design is sharp, there is no separate compliance margin; this column characterizes the local observations receiving the greatest weight in the estimate of τ. *PA 5–10k* is the same Pennsylvania 2020 cross-section, equal-weighted, restricted to municipalities with 5,000–10,000 residents: the broader sample that enters the regression before kernel weighting. *US 5–10k excl. PA* is the 2020 Census Gazetteer / ACS 2018–22 universe restricted to the same population × density window (population in $[5{,}000, 10{,}000)$, density in $[0, 300 + 845]$ ppsm) and to states other than Pennsylvania, weighted by 2020 population: the national portability target. Cross-column comparisons are discussed in the text. Median age is unavailable in the PA administrative panel (only age-bin shares); the US-band median (41.8) is reported for reference. 1,308 U.S. municipalities in the national comparison window, of which 180 are in Pennsylvania.

Table B3: Placebo cutoffs: per-cutoff diagnostics

c	N_L^{muni}	N_R^{muni}	h_{tot}	$\hat{\tau}_{\text{tot}}$	t_{tot}	$\hat{\tau}_{\text{res}}$	t_{res}	$\hat{\tau}_{\text{com}}$	t_{com}	sample
100	9	50	100	-0.284	-0.34	-0.878	-1.41	0.525	0.35	placebo
125	20	66	100	0.097	0.28	0.114	0.48	0.303	0.63	placebo
150	33	70	100	-0.222	-0.88	-0.251	-1.22	0.372	0.94	placebo
175	45	70	100	0.481	1.72	0.283	1.19	0.342	0.84	placebo
300	126	130	845	0.690	3.54	0.426	3.10	0.905	2.96	**true**
425	67	46	124	0.323	1.65	0.065	0.38	0.552	1.86	placebo
450	70	50	149	0.071	0.41	-0.102	-0.79	0.178	0.67	placebo
475	75	49	174	-0.186	-1.19	-0.086	-0.82	-0.352	-1.37	placebo
500	83	41	199	-0.512	-2.98	-0.159	-1.41	-0.880	-2.84	placebo
525	87	42	224	-0.353	-1.94	-0.236	-1.80	-0.448	-1.54	placebo
550	91	38	249	0.064	0.40	-0.111	-0.82	0.259	1.15	placebo
575	96	34	274	0.124	0.69	-0.357	-2.20	0.522	1.88	placebo
600	101	30	299	-0.099	-0.48	-0.233	-1.91	-0.037	-0.11	placebo
625	103	29	324	-0.207	-1.07	-0.041	-0.35	-0.376	-1.14	placebo
650	107	25	349	-0.042	-0.24	0.060	0.54	-0.200	-0.68	placebo
675	109	24	374	0.030	0.11	0.174	1.11	-0.129	-0.31	placebo
700	111	22	399	-0.107	-0.45	-0.016	-0.09	-0.216	-0.60	placebo

Notes: One row per cutoff in the placebo grid. N_L^{muni} and N_R^{muni} count unique municipalities with running variable below and above c within the bandwidth h_{tot} used for the total-recycling regression. $\hat{\beta}$ are PPML coefficients on the treatment indicator at c; t_{tot} is the z-statistic for total recycling. Specification held constant across cutoffs (see Figure 5 notes). The row at $c = 300$ is the true cutoff and uses the full baseline bandwidth window. Source: PA DEP administrative records, 2001–2024.

Table C1: Numerical decomposition of the reconstruction residual ε (log-points)

Outcome	ε	$\varepsilon_{\text{Jensen}}$ (exact)	ε_ω	ρ	$\max_{j,k} \lvert\Delta_j - \Delta_k\rvert$	2nd-order Jensen approx.
Total recycling	0.002	0.010	0.000	0.005	0.455	0.009
Residential recycling	−0.015	0.058	−0.001	−0.020	0.794	0.053
Commercial recycling	−0.043	0.059	0.002	−0.103	1.970	0.050

Notes: Material partition $\mathcal{K}_3 = \{\ell, m, n\}$ (leaf, non-leaf covered, non-covered). Column 2 reproduces the reconstruction residual $\varepsilon \equiv \tau - (\tau_a + \sum_k \tau^k)$ from Table C4. Column 3 reports the exact Jensen residual $\log(\sum_k \omega_k e^{\Delta_k}) - \sum_k \omega_k \Delta_k$, with $\Delta_k \equiv \tau_k^g + \tau_k^y$ the implied per-group total log-jump. Column 4 reports the weight-estimation wedge $\varepsilon_\omega = \sum_k (\bar{\omega}_k - \omega_k)\Delta_k$ from (14), using the bootstrap means $\bar{\omega}_k$ and the table weights ω_k. Column 5 reports $\rho = \varepsilon - \varepsilon_{\text{Jensen}} - \varepsilon_\omega$ from Equation (15). Column 6 reports the largest pairwise between-group log-jump gap $\max_{j,k} \lvert\Delta_j - \Delta_k\rvert$. Column 7 reports the second-order approximation $\frac{1}{2}\sum_{j<k} \omega_j \omega_k (\Delta_j - \Delta_k)^2$. All values in log-points; small discrepancies in the last digit reflect rounding.

Table C2: δ-sensitivity of Lee bounds on per-channel intensification τ_k^y.

	τ_ℓ^y		τ_m^y		τ_n^y	
δ	Lower	Upper	Lower	Upper	Lower	Upper
Panel A: Total recycling per capita						
0.000	-2.084	1.178	0.344	0.610	-0.372	0.598
0.010	-2.080	1.171	0.321	0.606	-0.434	0.585
0.025	-2.143	1.144	0.281	0.594	-0.517	0.572
0.050	-2.290	1.113	0.229	0.572	-0.667	0.550
Panel B: Residential recycling per capita						
0.000	-2.132	1.093	0.021	0.199	-0.949	0.691
0.010	-2.210	1.083	0.013	0.195	-0.988	0.681
0.025	-2.240	1.049	0.001	0.182	-1.031	0.660
0.050	-2.261	0.974	-0.018	0.162	-1.161	0.644
Panel C: Commercial recycling per capita						
0.000	-2.879	2.850	0.236	1.094	-0.099	0.523
0.010	-2.880	2.201	0.208	1.085	-0.183	0.518
0.025	-2.562	2.128	0.164	1.064	-0.299	0.507
0.050	-2.796	1.902	0.062	1.051	-0.461	0.493

Notes: Material partition $\mathcal{K}_3 = \{\ell, m, n\}$ (leaf, non-leaf covered, non-covered). Each panel reports Lee bound endpoints on the three intensification jumps τ_k^y, evaluated at four monotonicity-violation fractions δ across the three outcome definitions: total, residential, and commercial recycling. For a baseline trim share q_C^k, the lower endpoint uses $q_C^k + \delta$ and the upper endpoint uses $\max\{q_C^k - \delta, 0\}$. In the reported grid, $q_C^k + \delta < 1$ for every margin and outcome. Bounds are computed using the column-(3) specification of Table 2, the same specification used for the main decomposition in Table C4; the Lee construction is described in Appendix C.3. At $\delta = 0$ the table reproduces the Lee bounds on τ_k^y reported in the main decomposition table.

Table C3: Decomposition with 1980 pre-Act covariate adjustment

	Full sample		Curbside subsample	
	(1) No controls	(2) With controls	(3) No controls	(4) With controls
Panel A: Total recycling per capita				
Reduced form (τ)	0.637***	0.690***	0.556***	0.703***
	(0.191)	(0.195)	(0.182)	(0.202)
Activation (τ_a)	0.006	0.001	0.015*	0.014
	(0.010)	(0.011)	(0.009)	(0.009)
Composition ($\sum_k \omega_k \tau_k^g$)	0.095	0.096	0.093	0.098
Leaf composition	0.067	0.069	0.074	0.077
Covered composition	0.004	0.004	-0.005	-0.001
Non-covered composition	0.024	0.023	0.024	0.022
Intensification ($\sum_k \omega_k \tau_k^y$)	0.522	0.577	0.489	0.637
Leaf intensification	0.029	0.021	0.036	0.032
Covered intensification	0.309	0.319	0.273	0.300
Non-covered intensification	0.183	0.237	0.180	0.305
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.622	0.673	0.596	0.749
Curvature residual ε	0.015	0.016	-0.041	-0.047
Panel B: Residential recycling per capita				
Reduced form (τ)	0.436***	0.426***	0.334**	0.365**
	(0.144)	(0.137)	(0.151)	(0.145)
Activation (τ_a)	0.018	0.011	0.016	0.013
	(0.016)	(0.017)	(0.011)	(0.011)
Composition ($\sum_k \omega_k \tau_k^g$)	0.198	0.201	0.192	0.203
Leaf composition	0.145	0.147	0.141	0.148
Covered composition	0.012	0.012	-0.006	-0.003
Non-covered composition	0.040	0.041	0.057	0.057
Intensification ($\sum_k \omega_k \tau_k^y$)	0.183	0.179	0.138	0.179
Leaf intensification	0.040	0.027	0.051	0.050
Covered intensification	0.096	0.091	0.018	0.033
Non-covered intensification	0.047	0.061	0.069	0.096
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.399	0.391	0.346	0.395
Curvature residual ε	0.036	0.035	-0.013	-0.030

Table C3 (continued)

	Full sample		Curbside subsample	
	(1) No controls	(2) With controls	(3) No controls	(4) With controls
Panel A: Commercial recycling per capita				
Reduced form (τ)	0.795***	0.905***	0.735**	0.995***
	(0.287)	(0.305)	(0.287)	(0.332)
Activation (τ_a)	0.040	0.034	0.040	0.041
	(0.040)	(0.040)	(0.045)	(0.045)
Composition ($\sum_k \omega_k \tau_k^g$)	0.092	0.100	0.102	0.097
Leaf composition	0.010	0.009	0.006	0.006
Covered composition	0.058	0.066	0.057	0.059
Non-covered composition	0.024	0.025	0.038	0.033
Intensification ($\sum_k \omega_k \tau_k^y$)	0.704	0.802	0.623	0.907
Leaf intensification	0.008	0.008	0.001	0.003
Covered intensification	0.423	0.422	0.360	0.382
Non-covered intensification	0.273	0.372	0.262	0.522
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.836	0.936	0.764	1.045
Curvature residual ε	-0.042	-0.031	-0.030	-0.050

Notes: Decomposition on the material partition $\mathcal{K}_3 = \{\ell, m, n\}$ (leaf, non-leaf covered, non-covered) reported across four specifications: columns (1)–(2) use the full RD sample with and without the twelve 1980 pre-Act covariates entered linearly in the PPML index; columns (3)–(4) repeat the comparison on the baseline-curbside subsample. Row order matches Tables C4 and C6: reduced form τ, activation τ_a, the composition aggregate $\sum_k \omega_k \tau_k^g$ with its three constituents, the intensification aggregate $\sum_k \omega_k \tau_k^y$ with its three constituents, the reconstruction $\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$, and the curvature residual ε. Kernel (triangular), bandwidth ($h = 845$ ppsm), year and county fixed effects, baseline material-share weights ω_k, and cluster (municipality) match the main decomposition specification. The reduced-form and activation rows show the PPML treatment coefficient with municipality-clustered analytic standard errors in parentheses. The composition and intensification subrows, the reconstruction, and the curvature residual report point estimates only; Lee bounds for the analogous rows under the no-controls specifications are reported in Tables C4 and C6.

Table C4: Decomposing the mandate effect in the full sample: activation, composition, and intensification

	$\hat{\tau}$	SE	Lee bounds
Panel A: Total recycling per capita			
Reduced form (τ)	0.637^{***}	(0.213)	
Activation (τ_a)	0.006	(0.011)	
Composition ($\sum_k \omega_k \tau_k^g$)	0.095		**[0.09, 0.11]**†
Leaf composition	0.067		**[0.07, 0.07]**
Covered composition	0.004		**[0.00, 0.01]**
Non-covered composition	0.024		**[0.02, 0.03]**
Intensification ($\sum_k \omega_k \tau_k^y$)	0.522		[-0.14, 0.66]†
Leaf intensification	0.029		[-0.19, 0.11]
Covered intensification	0.309		**[0.19, 0.33]**
Non-covered intensification	0.183		[-0.13, 0.22]
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.622^{***}	(0.210)	
Curvature residual ε	0.015		
Panel B: Residential recycling per capita			
Reduced form (τ)	0.436^{***}	(0.164)	
Activation (τ_a)	0.018	(0.018)	
Composition ($\sum_k \omega_k \tau_k^g$)	0.198		**[0.19, 0.23]**†
Leaf composition	0.145		**[0.14, 0.16]**
Covered composition	0.012		**[0.01, 0.02]**
Non-covered composition	0.040		**[0.03, 0.05]**
Intensification ($\sum_k \omega_k \tau_k^y$)	0.183		[-0.53, 0.44]†
Leaf intensification	0.040		[-0.41, 0.21]
Covered intensification	0.096		**[0.01, 0.13]**
Non-covered intensification	0.047		[-0.13, 0.09]
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.399^{***}	(0.153)	
Curvature residual ε	0.036		

(Continued on next page; full notes follow Table C4.)

Table C4: (continued from previous page)

	$\hat{\tau}$	SE	Lee bounds
Panel A: Commercial recycling per capita			
Reduced form (τ)	0.795**	(0.318)	
Activation (τ_a)	0.040	(0.045)	
Composition ($\sum_k \omega_k \tau_k^g$)	0.092		**[0.07, 0.16]**†
Leaf composition	0.010		[0.00, 0.01]
Covered composition	0.058		**[0.04, 0.11]**
Non-covered composition	0.024		**[0.02, 0.04]**
Intensification ($\sum_k \omega_k \tau_k^y$)	0.704		**[0.03, 0.79]**†
Leaf intensification	0.008		[-0.02, 0.02]
Covered intensification	0.423		**[0.10, 0.48]**
Non-covered intensification	0.273		[-0.05, 0.29]
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.836**	(0.330)	
Curvature residual ε	-0.042		

Notes: Three-group nested decomposition of the mandate-induced cutoff jump τ by recycling outcome $S \in \{\text{tot}, \text{resid}, \text{comm}\}$ and material group $k \in \mathcal{K}_3 = \{\ell, m, n\}$, where ℓ is leaf waste, m is non-leaf covered materials, and n is non-covered tonnage. Identity: $\tau \approx \tau_a + \sum_{k \in \mathcal{K}_3} \omega_k(\tau_k^g + \tau_k^y) + \varepsilon$, where τ_a is the activation margin, τ_k^g is material-group appearance (the composition margin), and τ_k^y is within-group tonnage intensification. Material groups: *leaf* = DEP column Y01 ("Yard and Leaf Waste"); *non-leaf covered* = paper, glass, metal, eligible plastics, single-stream, commingled, and aggregated paper/cardboard categories assigned to the covered group in Section 6; *non-covered* = all remaining non-statutory tonnage. Weights ω_k are control-side, kernel-weighted shares of per capita tonnage, so $\omega_c = \omega_\ell + \omega_m$. Each panel reports the reduced-form $\hat{\tau}$ and then regroups the identity two ways: by margin of adjustment and by material-group partition. In the by-margin footer, Composition denotes $\sum_{k \in \mathcal{K}_3} \omega_k \tau_k^g$ and Intensification denotes $\sum_{k \in \mathcal{K}_3} \omega_k \tau_k^y$. In the material-group footer, the contribution of group k is $\tau^k = \omega_k(\tau_k^g + \tau_k^y)$. The reduced-form row uses analytic delta-method standard errors. All other point-identified rows report bootstrap standard errors ($B = 1{,}000$, clustered by municipality). Set-identified rows report Lee (2009) bounds and carry no stars. Lee bounds in **bold** indicate intervals that strictly exclude zero. Bounds marked † are partition-aggregated Lee bounds: endpoints sum low and high endpoints separately across the constituent material groups. All estimates use the specification of Table 2, column (3), matching the main decomposition. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$ on point-identified rows.

Table C5: Baseline (2001) curbside status at the 300 ppsm cutoff

	(1) No controls	(2) 1980 controls
RD estimate at 300 ppsm	0.290	0.257
	(0.088)	(0.084)
Robust p-value	0.303	0.485
Observations	182	182
Bandwidth (ppsm)	845	845

Notes: Sharp RD estimate of the discontinuity in 2001 curbside operation at the 300 ppsm cutoff, estimated on the cleaned 2001 cross-section used to define the baseline-curbside subsample. Triangular kernel and bandwidth $h = 845$ ppsm, matching the main analysis. Conventional point estimates with heteroskedasticity-robust standard errors in parentheses; reported p-values are robust bias-corrected p-values from Calonico et al. (2014). Column (2) includes the twelve 1980 pre-Act controls used in the benchmark specification. Source: PA DEP curbside registry, 2001 cross-section.

Table C6: Decomposing the mandate effect in the curbside subsample: activation, composition, and intensification

	$\hat{\tau}$	SE	Lee bounds
Panel A: Total recycling per capita (curbside subsample)			
Reduced form (τ)	0.556***	(0.182)	
Activation (τ_a)	0.015*	(0.009)	
Composition ($\sum_k \omega_k \tau_k^g$)	0.093		**[0.09, 0.10]**†
Leaf composition	0.074		**[0.07, 0.07]**
Covered composition	-0.005		**[-0.00, -0.00]**
Non-covered composition	0.024		**[0.02, 0.03]**
Intensification ($\sum_k \omega_k \tau_k^y$)	0.489		[-0.13, 0.63]†
Leaf intensification	0.036		[-0.18, 0.12]
Covered intensification	0.273		**[0.18, 0.29]**
Non-covered intensification	0.180		[-0.13, 0.22]
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.596***	(0.182)	
Curvature residual ε	-0.041		
Panel B: Residential recycling per capita (curbside subsample)			
Reduced form (τ)	0.334**	(0.151)	
Activation (τ_a)	0.016	(0.011)	
Composition ($\sum_k \omega_k \tau_k^g$)	0.192		**[0.19, 0.20]**†
Leaf composition	0.141		**[0.14, 0.15]**
Covered composition	-0.006		**[-0.01, -0.00]**
Non-covered composition	0.057		**[0.06, 0.06]**
Intensification ($\sum_k \omega_k \tau_k^y$)	0.138		[-0.51, 0.41]†
Leaf intensification	0.051		[-0.34, 0.21]
Covered intensification	0.018		[-0.05, 0.04]
Non-covered intensification	0.069		[-0.11, 0.16]
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.346**	(0.151)	
Curvature residual ε	-0.013		

(Continued on next page; full notes follow Table C6.)

Table C6: (continued from previous page)

	$\hat{\tau}$	SE	Lee bounds
Panel A: Commercial recycling per capita (curbside subsample)			
Reduced form (τ)	0.735**	(0.287)	
Activation (τ_a)	0.040	(0.045)	
Composition ($\sum_k \omega_k \tau_k^g$)	0.102		**[0.08, 0.16]**†
Leaf composition	0.006		[-0.00, 0.01]
Covered composition	0.057		**[0.05, 0.10]**
Non-covered composition	0.038		**[0.04, 0.05]**
Intensification ($\sum_k \omega_k \tau_k^y$)	0.623		**[0.03, 0.69]**†
Leaf intensification	0.001		[-0.01, 0.01]
Covered intensification	0.360		**[0.14, 0.40]**
Non-covered intensification	0.262		[-0.10, 0.28]
Reconstruction ($\tau_a + \sum_k \omega_k(\tau_k^g + \tau_k^y)$)	0.764***	(0.287)	
Curvature residual ε	-0.030		

Notes: Decomposition restricted to the 2001-curbside subsample (`has_cs_0imp`=1). All other modeling choices match Table C4: PPML separate slopes, triangular kernel with $h = 845$ ppsm, dep-year and county fixed effects, twelve 1980 pre-Act controls, cluster (municipality) standard errors. Material groups: ℓ leaf (DEP Y01), m non-leaf covered (paper, glass, metal, plastic, single-stream, commingled, bundled), n non-covered. Weights ω_k are control-side kernel-weighted shares of per-capita tonnage on the curbside subsample. The reduced-form and reconstruction rows use municipality-clustered analytic standard errors from the PPML estimation. Bootstrap is omitted on this check, so standard errors are blank on the activation, composition, intensification, and per-channel contribution rows. Set-identified rows report Lee (2009) bounds; intervals in **bold** strictly exclude zero. Bounds marked † are partition-aggregated (lower and upper endpoints summed separately across constituent material groups). * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$ on point-identified rows.